Optimisation of Mutation Detection in Genes Responsible for Seed Shattering and Seed Size in Perennial Ryegrass (*Lolium perenne* L.)

A thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy

in Cellular and Molecular Biology

School of Biological Science

University of Canterbury

by

Zeyu Fu

2017

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Deputy Vice-Chancellor's Office
Postgraduate Office

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 2 has been published: Fu Z-Y., Song J-C. and Jameson PE. (2017) *A rapid and cost effective protocol for plant genomic DNA isolation using regenerated silica columns in combination with CTAB extraction.* Journal of Integrative Agriculture 16(8): 1682-1688. http://dx.doi.org/10.1016/S2095-3119(16)61534-4. (Journal Articles)

Please detail the nature and extent (%) of contribution by the candidate:

Zeyu Fu (75%) and Jiancheng Song (25%) designed the experiments. Zeyu Fu carried out the experiments and analysed the data.

Zeyu Fu wrote the manuscript (85%) with significant input from Prof. Paula Jameson (15%).

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: Paula Jameson

Signature: Paula E Jameson

Date: 15/10/2017

Table of Contents

Acknowledgements **1**

Abstract **2**

1 Chapter 1: Introduction **4**

1.1 Background **4**

1.2 Molecular mechanisms of plant organ abscission **7**

1.3 Genetic mechanism of the shattering trait in monocots **9**

1.4 Cytokinins and plant productivity **17**

1.5 Reverse genetics and TILLING **21**

1.6 Aims and objectives **22**

2 Chapter 2: DNA extraction method **24**

2.1 Introduction **24**

2.2 Materials and Methods **25**

2.2.1 Column regeneration 25

2.2.2 DNA extraction procedure 25

2.2.3 Yield and quality determination 26

2.2.4 Assessing DNA quality using restriction enzymes 27

2.2.5 Quality check using amplification of genomic DNA by RT-qPCR 27

2.2.6 Testing for DNA cross contamination 27

2.3 Results and discussion **28**

2.3.1 Modification of the CTAB method 28

2.3.2 Column regeneration 29

2.3.3 Yield and quality of DNA were not compromised when regenerated columns were used 30

2.3.4 Sample-to-sample cross contamination did not occur 33

2.4 Conclusions **34**

3 Chapter 3: Identification and expression of putative seed shattering genes in *Lolium perenne* L. **36**

3.1 Introduction **36**

3.2 Material and methods **38**

3.2.1 Plant material 38

3.2.2 RNA isolation and cDNA synthesis 42

3.2.3 Candidate seed shattering genes in perennial ryegrass 43

3.2.4 Primer design and optimisation for gene expression study 46

3.2.5 Reference gene identification 47
3.2.6 Target gene sequence verification 48
3.2.7 Reverse transcription quantitative PCR (RT-qPCR) 48
3.2.8 Cloning of the full length *LpSH1* gene 50
3.3 Results 52
3.3.1 RNA extraction and cDNA synthesis 52
3.3.2 Primer pair selection and PCR optimisation for gene expression study 53
3.3.3 Isolation of candidate seed shattering genes from perennial ryegrass 55
3.3.4 Relative expression of candidate seed shattering genes in perennial ryegrass 57
3.3.5 Determination of the full length *LpSH1* gene 65
3.4 Discussion 68
4 Chapter 4: Optimisation of method for mutation screening 72
4.1 Introduction 72
4.2 Material and methods 75
4.2.1 Development of a method combining HRM analysis with CEL I nuclease to screen for mutations 75
4.2.2 Testing the CEL I+HRM method using *LpCKX1* 81
4.3 Results 84
4.3.1 HRM analysis in combination with CEL I 84
4.3.2 Sequencing of full length of *LpCKX1* and part of *LpCKX2* 94
4.3.3 Detection of CEL I digestion products from *LpCKX1* 95
4.4 Discussion 97
4.4.1 CEL I digestion with HRM analysis 97
4.4.2 Sequence determination of *LpCKX1* and *LpCKX2* genes 99
4.4.3 Detection of mutations in *LpCKX1* using the optimised CEL I+HRM method 100
4.4.4 Summary 100
5 Chapter 5: Mutation screening 101
5.1 Introduction 101
5.2 Material and methods 102
5.2.1 Optimisation of homemade HRM master mixes 102
5.2.2 Plant materials and DNA extraction 104
5.2.3 First run of nested PCR for mutation screening of *LpSH1* and *LpCKX1* 105
5.2.4 Second run PCR and HRM analysis 106
5.2.5 Sequencing and sequence analysis 108
5.3 Results 109
5.3.1 Trial of homemade HRM master mix 109

5.3.2 Mutation screening in *LpSH1* 111
5.3.3 Mutation screening in *LpCKX1* 112
5.3.4 Sequencing of *LpSH1* and *LpCKX1* 113
5.4 Discussion 120
5.4.1 Trial of a variety of HRM master mix 120
5.4.2 Mutation screening 121
6 Chapter 6: Final discussion 125
6.1 A genetic model for seed shattering in perennial ryegrass 125
6.2 Cytokinin oxidase/dehydrogenase 127
6.3 Mutation screening 128
6.4 Future study 130
References 133
Appendices 147

LIST OF FIGURES

Figure 1.1 Abscission layer morphology in rice (A-D) (Zhou et al., 2012) and ryegrass (E-I) (Elgersma et al., 1988). 6
Figure 1.2 Inferred evolutionary history of two shattering loci in two sorghum species (Lin et al., 2012).. 15
Figure 1.3 A Genetic Model of Regulatory Network Specifying abscission layer Development in Rice (Zhou et al., 2012).....16
Figure 1.4 The most recent genetic model for the abscission process in rice (Yoon et al., 2017). 17
Figure 1.5 Scheme of CK biosynthesis, interconversions and degradation in plants. Enzymes involved in these interconversions are indicated by red numbers. 19
Figure 2.1 Modification of CTAB method to reduce lysis incubation time. 29
Figure 2.2 Optimization of the regeneration protocol. 30
Figure 2.3 Comparison of yield and integrity of DNA obtained from a new column and columns regenerated up to 11 times (Re1–11). 31
Figure 2.4 Quality assessment using restriction enzyme assay of DNA eluted from a new column and columns regenerated up to 11 times (Re1–11).....32
Figure 2.5 Amplification of genomic DNA with RT-qPCR.....33
Figure 2.6 Comparison of PCR products after elution of pea genomic DNA from new and regenerated columns. 34
Figure 3.1 Spike and spikelet of ryegrass. 37
Figure 3.2 The early developmental stages of spikes.. 39
Figure 3.3 RNA integrity check.....53
Figure 3.4 The RT-qPCR performance with different ryegrass reference gene primer pairs..54
Figure 3.5 Rooted phylogenetic tree of candidate target seed shattering homologues in monocots.. 56
Figure 3.6 Relative expression of putative *LpqSH1*, *LpSH5*, *LpSH4*, *LpSHAT1*, *LpSH1* *LpLG1*, *LpWRKY*, and *LpQ* in perennial ryegrass cv. RI009 (first biological replicate). The plant materials were collected the glasshouse at the University of Canterbury, New Zealand, from October to December in 2014. 59
Figure 3.7 Relative expression of putative *LpqSH1*, *LpSH5*, *LpSH4*, *LpSHAT1*, *LpSH1* *LpLG1*, *LpWRKY*, and *LpQ* in perennial ryegrass cv. Nui (second biological replicate).

The plant materials were collected from a field located outside of Christchurch, New Zealand, from October to December in 2014. 60
Figure 3.8 Relative expression of putative *LpqSH1*, *LpSH5*, *LpSH4*, *LpSHAT1*, *LpSH1* *LpLG1*, *LpWRKY*, and *LpQ* in perennial ryegrass cv. Nui (third biological replicate). The plant materials were collected from the glasshouse in Yantai University, China, from May to June in 2015. 61
Figure 3.9 The overall pattern of relative expression level of candidate seed shattering genes in first biological replicate. The details are described in Figure 3.6. 62
Figure 3.10 The overall pattern of relative expression level of candidate seed shattering genes in second biological replicate. The details are described in Figure 3.7. 63
Figure 3.11 The overall pattern of relative expression level of candidate seed shattering genes in third biological replicate. The details are described in Figure 3.8. 64
Figure 3.12 Gene structure of *LpSH1* 65
Figure 3.13 Amino acid sequence alignment of LpSH1 and homologous gene proteins. 66
Figure 3.14 Rooted phylogenetic tree of the YABBY2 sub-family in monocots. 67
Figure 4.1 Raw channel (A) and melt curve (B) of HRM analysis. The dissociation temperature is indicated by the arrow. 73
Figure 4.2 SDS protein gel of CEL I. 85
Figure 4.3 The PCR products of GC-CELI-F1/R1, F2/R1, F3/R2, F4/R3, F/R373, F/R648, F/R902 amplified with G/C plasmids from left to right, respectively (refer to Table 4.1 for sequences). 85
Figure 4.4 The cleavage products from three different master mixes.. 87
Figure 4.5 CEL I digestion products with different treatment refer to Table 4.3. 88
Figure 4.6 Melting curve in HRM analysis of CEL I digestion products with different treatments, refer to Table 4.3. 88
Figure 4.7 The PCR products from GC-CELI-F/R373 and F/R648 with different treatments in EvaGreen and template amount for CEL I digestion. 90
Figure 4.8 The melting curves of CEL I digestion of PCR products from GC-CELI-F/R373 with different Treatments. 90
Figure 4.9 The ten PCR products with different treatments. 91
Figure 4.10 The melting curve of ten digestion reactions with different treatments as described in Table 4.5 and Figure 4.9. The digested band peaks are indicated by the arrow. 92
Figure 4.11 The digestion products from the different ratios of mixed PCR products. 93

Figure 4.12 The melting curves of digestion samples from the different ratios of digestion template. 93
Figure 4.13 Phylogenetic tree of *CKX1* (A) and *CKX2* (B) sub-family in monocots. 94
Figure 4.14 Gene structures of *LpCKX1* and *LpCKX2* and the approximate sites of the primers. 95
Figure 4.15 Cleavage products of exon 1 in *LpCKX1* from cv. Nui in Trial 1 (A) and Trial 2 (B). 96
Figure 5.1 Structures of the gene region to screen mutations in (A) *LpSH1* and (B) *LpCKX1*, and primers location at approximate primers sites. 107
Figure 5.2 HRM normalised fluorescence graphs of GC-HRM-F1/R1 performed by the Biotium HRM Kit (A) and AmpliTaq homemade master mix (B). 110
Figure 5.3 HRM normalised fluorescence graphs of Allele4/8-HRM-F/R performed by the Agilent HRM Kit (A) and AmpliTaq homemade master mix (B). 111
Figure 5.4 The normalised fluorescence (A, B, C) and melting curves (D, E, F) of LpCKX1-HRM-F4R5 from three diploid ryegrass cultivars, including cv. Nui (A and D), cv. RI009 (B and E); and cv. Stellar (C and F) 113
Figure 5.5 Alignment of sequences from mutation screening region in *LpSH1*. The SNPs and INDEL are highlighted in blue, and the exons and introns are indicated by red arrows. 115
Figure 5.6 The alignment of the coding region of *LpSH1*. Seven SNPs are highlighted in blue, and the exons 4 and 5 are underlined by red and blue. 116
Figure 5.7 The alignment of translated amino acid sequences from exons 4 and 5 in *LpSH1*. 116
Figure 5.8 Alignment of sequences from mutation screening region in *LpCKX1*. The SNPs and INDELs are highlighted in blue, and the exons 2 and 3 and intron 2 are indicated by red arrows 117
Figure 5.9 Multiple alignment of part of the coding region of *LpCKX1*.. 118
Figure 5.10 Multiple alignment of translated amino acid sequences from exons 2 and 3 in *LpCKX1*. 119
Figure 5.11 Details of the SNPs in (A) the G/C plasmid and (B) the allele 4/8. 121
Figure 5.12 The flow chat of the mutation screening process used in this project. 124
Figure 6.1 Genetic model for seed shattering in perennial ryegrass. 127

LIST OF TABLES

Table 2.1 Primers sequences.....27
Table 3.1 Plant material collected for the first biological replicate for RNA extraction. Material at each stage was collected from different individuals.....41
Table 3.2 Plant material collected for the second biological replicate for RNA extraction. Material at each stage was collected from different individuals.....41
Table 3.3 Plant material collected for the third biological replicate for RNA extraction. Material at each stage was collected from different individuals.....42
Table 3.4 The sources of homologous genes described in this thesis.....44
Table 3.5 Primer sequences for gene expression study 47
Table 3.6 The components of the 2x qPCR buffer 49
Table 3.7 The components of 2x homemade SYBR Green master mix.....49
Table 3.8 Reaction composition using 2x SYBR Green homemade master mix 49
Table 3.9 Selected primer sequences for introns of *LpSH1* 51
Table 3.10 Reference and target gene primer sequences selected for the gene expression study.....54
Table 4.1 G/C plasmid primer sequences for CEL I digestion optimisation.....77
Table 4.2 The optimisation of CEL I digestion with different amounts of CEL I and reaction time. 78
Table 4.3 CEL I digestion with different treatment of two PCR products. 78
Table 4.4 Assessments of CEL I digestion using PCR products from G/C plasmid. 79
Table 4.5 Treatments of 10 CEL I digestion reactions 80
Table 4.6 Ratios of G and C plasmid PCR products in the CEL I digestion reaction 80
Table 4.7 Primer sequences for *LpCKX1* and *LpCKX2* 82
Table 4.8 Components of two trails for detection of CEL I+HRM method.....84
Table 5.1 Details of the PCR products of the two primer pairs. The SNP of each plasmid pair is highlighted in yellow and underlined. 103
Table 5.2 Primer sequences for G/C plasmid and Allele4/8.....103
Table 5.3 Details of each commercial HRM kit and DNA polymerases. 104
Table 5.4 Details of each cultivar and genomic DNA sample number. 105
Table 5.5 Primer sequences of *LpSH1* and *LpCKX1* for the first round of nested PCR.....106
Table 5.6 Components of the first run PCR master mixes for *LpSH1* and *LpCKX1*. 106

Table 5.7 Primer sequences for mutation screening in *LpSH1* and *LpCKX1* for HRM analysis. 108
Table 5.8 Individuals selected for sequencing in *LpSH1* and *LpCKX1* based on genotype. .109
Table 5.9 Pattern number for each primer pair following HRM analysis 112

Acknowledgements

It is my great pleasure to express my thanks to my supervisor Professor Paula Jameson for her advice, encouragement and help during my PhD study. I am very thankful to her for the valuable suggestions and careful correction of this work. I would like to thank Professor Jiancheng Song, my co-supervisor, for his continued support and valuable suggestions in my study.

I would like to thank my lab mates for their help over the past five years: Annu Ninan, Matthew Van Voorthuizen, Anish Shah, Samarth, Matthew Arnet, Qianqian Guo, Jessica Roach, Pragathi Dhandapani, and also Professor Matthew Turnbull for his support. Many thanks to lab manager Thomas Evans and glasshouse manager Dave Conder. I would also like to thank staff and students in Professor Song's lab at Yantai University for their help in the transgenic experiments and sample collection.

I gratefully acknowledge the Foundation for Arable Research for providing the financial support for this research. I would also like to thank Dr Alan Stewart, PGG Wrightson, for the supply of several seed lines.

Finally, thanks to all of my family and friends for their support and encouragement during my study and life in New Zealand.

Abstract

Perennial ryegrass (*Lolium perenne* L.) is one of the most important pasture grasses in New Zealand. However, seed production is negatively impacted by seed shattering (shedding). Recently, genes involved in the shattering process have been isolated and functionally characterised in several crop species, including *qSH1* and *SH4* in rice and *SH1* in sorghum. The aim of this project was to identify the critical genes involved in the seed shattering process in perennial ryegrass, and then to screen for mutations in the target gene in a perennial ryegrass EcoTILLING population. Additionally, seed size is an aspect of seed production of interest to perennial ryegrass breeders. *CKX1* has been identified as playing a key role in seed size in rice and wheat, so the *LpCKX1* was added as a target gene for mutation screening.

DNA sequences of genes involved in seed shattering and seed size in the Poaceae were used to identify and isolate target genes in perennial ryegrass using a comparative genomics strategy. The candidate ryegrass shattering genes were identified using an 'in-house' perennial ryegrass transcriptome database. The relative expression levels of candidate ryegrass shattering genes were determined using RT-qPCR during the floret and seed developmental stages. A genetic model for seed shattering in perennial ryegrass is outlined. *LpSH1* and *LpCKX1* were selected as the target genes for mutation screening in the ryegrass EcoTILLING population, and the full-length DNA sequences of both target genes were amplified.

Several methods designed to optimise the EcoTILLING approach were developed. For mutation screening, plant genomic DNA was isolated using regenerated silica columns. The reliability of using regenerated silica columns was tested in terms of both yield and purity. A new method combining high resolution melting (HRM) analysis to detect CEL I digestion products was developed to reduce the time taken for mutation screening. The CEL I+HRM method was tested to screen mutations in exon 1 of *LpCKX1*, but due to the high degree of genetic diversity in perennial ryegrass and the high GC-content in *LpCKX1*, this method was not suitable for perennial ryegrass, but could be used for self-pollinated species, such as wheat and pea. A homemade HRM master mix was optimised, and HRM analysis was used to screen for mutations in *LpSH1* and *LpCKX1* in the EcoTILLING population. Multiple mutations in *LpSH1* and *LpCKX1* were identified from the EcoTILLING population. These

were then sequenced to detect if the mutations were likely to cause a change in protein structure and function.

The strategy adopted in this project allowed for the application of genetic knowledge from well-studied domesticated plants to be applied to a lesser-studied crop plant, and shows the potential for detecting useful mutations for future plant breeding.

1 Chapter 1: Introduction

1.1 Background

Various ryegrass species are used worldwide as forage grasses, including *Lolium perenne* (perennial ryegrass) and *Lolium multiflorum* (Italian ryegrass). According to a report from MPI released in 2016, titled " How valuable is that plant species?", ryegrasses are the most valuable plant species in New Zealand, with a $14.5 billion impact on New Zealand's GDP in 2012 (MPI, 2016). Perennial ryegrass is highly valued for forage for the livestock industry, so the supply of high-quality perennial ryegrass seed is a significant issue for the agronomic sector. Around 150,000 ha is used by the New Zealand arable industry, of which around 50,000 ha is used for seed production from a range of species (Hampton et al., 2012). In 2011, about 14,000 ha was used for ryegrass seed production in New Zealand, with 54.3% of the area for perennial ryegrass (Hampton et al., 2012). Total New Zealand seed production was around 122,000 tonnes in 2010 and around 136,000 tonnes in 2011, of which ryegrass is the largest seed yield by volume, about 30,000 tonnes per annum (Hampton et al., 2012). Exports of ryegrass seed were valued at $38.96 million and $50.22 million in 2010 and 2011, respectively (Hampton et al., 2012).

However, there are a number of factors affecting seed yield in this species, including poor pollination and seed set, limited knowledge about nitrogen and water use efficiency, the effect of biotic and abiotic stresses, and large harvest losses (Charlesworth, 1989; Elgersma, 1990; Hampton et al., 2012). The primary cause of yield loss is seed shattering (i.e. the shedding of seed), which negatively impacts seed production during the harvest process. This trait may lead to a seed yield loss between 12 and 39% of the potential production of some seed crops, and the seed loss in grasses maybe as high as 75% (Simon, 1997). Therefore, seed shattering can result in considerable economic losses. In addition to seed loss, the seed shattering trait can lead to contamination of other arable crops. Cropping farmers use a crop rotation system on their farms, such as wheat, oat and ryegrass (Liebman & Dyck, 1993). Severe seed shattering in ryegrass negatively affects the next rotation causing seed contamination in the next year's crop.

Seed shattering is due to the abscission process in plants, a natural phenomenon and an adaptive trait. It is an important trait for seed dispersal in wild plant species but has been lost progressively during domestication (Purugganan, 2014). Genes preventing seed shattering

gradually became dominant in the early progenitors of crops, such as rice, wheat and sorghum (Konishi et al., 2006; Li et al., 2006; Simons et al., 2006; Lin et al., 2012; Purugganan, 2014). In the paper by Zhou et al. (2012), the abscission layer (AL) in wild rice is visible in the basal area adjacent to the sterile lemmas, and consists of several cell layers (Figure 1.1 A, C). The abscission process takes place between cell layers of the AL. In contrast, the AL in the non-shattering rice is completely absent (Figure 1.1 B, D).

In modern crops, non-shattering, homozygous recessive alleles substituted for the shattering trait resulting from unconscious selection over millennia (Lin et al., 2012). However, ryegrass is much less domesticated than modern crops, so the seed shattering trait is still retained. Based on Elgersma et al. (1988), perennial ryegrass also has an AL in a similar place to rice (Figure 1.1 E-G). The separation of the seed from the rachilla occurs at the AL, commencing from the outer sides of the rachilla (Figure 1.1 H, I)

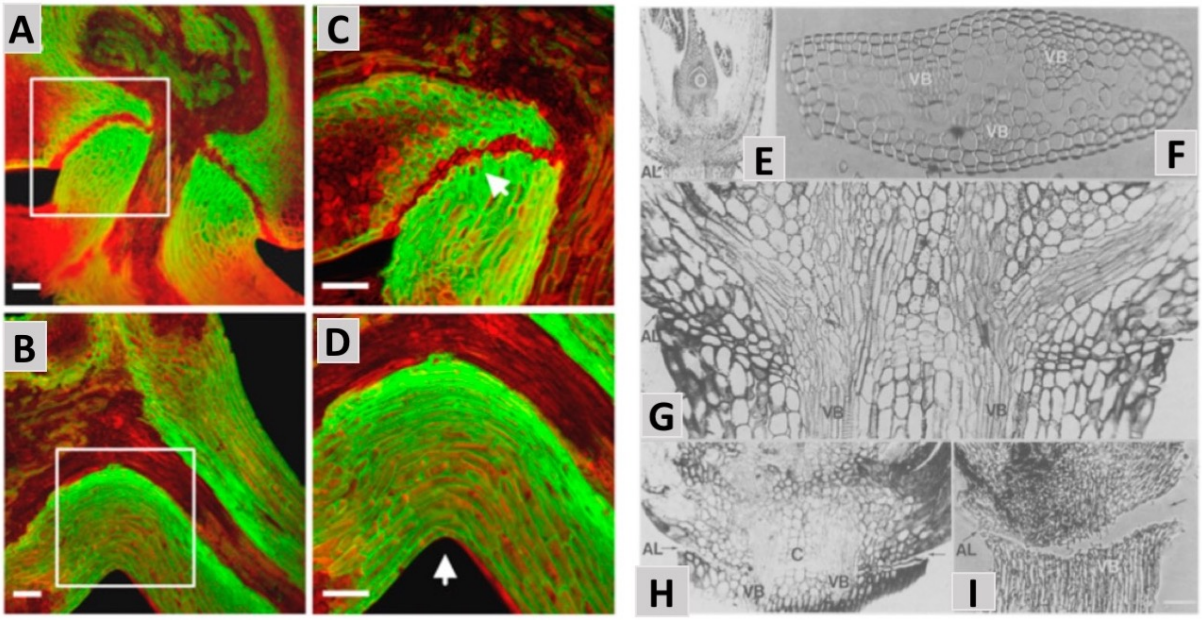


Figure 1.1 Abscission layer morphology in rice (A-D) (Zhou et al., 2012) and ryegrass (E-I) (Elgersma et al., 1988). A and B show fluorescence images of longitudinal sections across the flower and pedicel junction of rice stained by acridine orange. The white boxes indicate the region where the AL is located. Bars=50 µm. C and D show enlarged views corresponding to the white boxes in A and B, respectively. Arrows points to the AL in the wild rice and the corresponding region in a mutant line of rice. E show fluorescence images of longitudinal sections or perennial ryegrass. The AL is indicated by an arrow. F illustrates a cross section through the rachilla, showing three vascular bundles. G shows the AL at two weeks after heading. H shows the opening of the AL at five weeks after anthesis, starting from the epidermal sides. I shows the seed is completely separated at the area of the AL and the vascular bundles.

In addition to seed shattering, seed size is also an issue for perennial ryegrass seed production. If the seed number or weight can be increased, ryegrass seed production will be significantly increased, as communicated by the Foundation for Arable Research (FAR), who supplied funding for this project.

1.2 Molecular mechanisms of plant organ abscission

Leaves, flowers, fruits and other plant organs abscise from their mother plants when they become senescent. This developmentally programmed abscission happens within a specific tissue, the abscission layer (AL), or abscission zone (AZ). The position of the AL is classified into two types: one is at the boundary between organs, such as the base of floral organs or the base of the grain; the second is within an organ, for instance in the flower pedicel of tomato (Estornell et al., 2013).

An accepted model of abscission suggests that there are four major stages in the development and activation of the AL: (a) differentiation of the AL at the future site of organ detachment; (b) acquisition by the AL cells of competence to react to abscission signals (phytohormone regulation); (c) initiation and activation of the abscission process within the AL and organ detachment; (d) differentiation of a protective layer on the surface of the separation layer at the side of main plant body (Patterson, 2001).

The first step in the abscission process is related to the differentiation of a functionally specialised tissue that allows abscission. The differentiation is induced by several transcription factors (TFs), belonging to different gene families involved in the ontogeny of the AL.

In *Arabidopsis thaliana* (arabidopsis), the AL of flowers forms at the base of floral organs, such as petals or stamens. When flowers become senescent, abscission processes are activated in the AL. However, a plant with a double mutation in *BLADE-ON-PETIOLE1* (*BOP1*) and *BLADE-ON-PETIOLE2* (*BOP2*) never sheds the sepals, petals or stamens, even after senescence and wilting, because the mutation obstructs the differentiation of the AL(McKim et al., 2008). They also indicated that *BOP1/2* redundantly control the differentiation of AL of floral organs (McKim et al., 2008). In addition, the differentiation of the AL in seeds depends on the MADS-box TF gene *SEEDSTICK* (*STK*) and the bHLH family TF gene *HECATE3* (*HEC3*) (Pinyopich et al., 2003; Ogawa et al., 2009).

In tomato, the AL is formed in the pedicel region, with a knuckle or joint-like appearance. Two genes, *JOINTLESS* (*J*) and *JOINTLESS2* (*J2*), are involved in the AL formation. The tomato containing a double mutation of *j* and *j2* failed to differentiate an AL and then developed a pedicel without an AL (Mao et al., 2000). Nakano et al. (2012) suggested that

MACROCALYX (*MC*), belonging to the *APETALA1* family of MADS-box genes, plays a role in the formation of the pedicel AL in tomato and that MC and JOINTLESS could comprise a heterodimer binding to a DNA motif, which means that MC adjusts pedicel AL differentiation by forming a MADS-box complex with JOINTLESS.

Recently, several genes relating to the differentiation and formation of the pedicel AL have been identified in rice, for example, the MYB3 family TF gene *OsSH4* (Li et al., 2006); the BEL1 family TF gene *OsqSH1* (Konishi et al., 2006) and *OsSH5* (Yoon et al., 2014); the AP2 family TF gene *SHATTERING ABORTION 1* (*OsSHAT1*) (Zhou et al., 2012); the YABBY family TF gene, *OsSH1*, which is a homologue of the sorghum (*Sorghum bicolor*) *SHATTERING 1* (*SH1*) (Lin et al., 2012); and *OsSH-H* encoding carboxy-terminal domain phosphatase-like 1 (OsCPL1) protein (Ji et al., 2010). Of them, *OsSH4, OsqSH1, OsSHAT1, OsSH1* positively regulate the pedicel AL formation, but *SH-H* represses the differentiation of the AL. Monocotyledon seed shattering genes are covered in more detail in Section 1.3.

Once the AL is properly differentiated, its cells acquire competence to respond to developmental and environmental factors that trigger abscission (Estornell et al., 2013; Nakano & Ito, 2013). The response of AL cells to abscission triggering signals is reported to be mediated by phytohormone regulation through modulating the expression of abscission-related genes during the activation phase of abscission (Meir et al., 2006; Meir et al., 2010).

Several plant hormones, including ethylene, abscisic acid (ABA), jasmonic acid (JA), and in specific circumstances, cytokinin, positively affect the abscission process, as abscission-accelerating signals; while others, such as auxin, gibberellins, polyamines and brassinosteroids play negative roles, as abscission-inhibiting signals (Estornell et al., 2013). The interplay between ethylene signalling and auxin responses mainly controls activation of the abscission process, primarily determining the timing of abscission (Meir et al., 2010).

Ethylene plays an important role as a positive regulator of abscission. Ethylene treatment can increase abscission of leaves, flowers and fruits; conversely, treatment with inhibitors of ethylene biosynthesis, such as amino-ethoxyvinylglycine (AVG), can block abscission (Brown, 1997). Ethylene induces expression of genes encoding cell wall remodelling enzymes and their secretion to the cell wall (Brown, 1997). However, in arabidopsis, mutations in the ethylene receptor *ETHYLENE RESPONSE1* (*ETR1*) and ethylene signalling gene *ETHYLENE-INSENSITIVE2* (*EIN2*) suggested that ethylene probably just plays an

assistant role in abscission (Patterson & Bleecker, 2004). In arabidopsis, an alternative pathway to regulate organ abscission without ethylene has also been proposed, which means that ethylene may not be essential for abscission activation (Bleecker & Patterson, 1997).

While application of ethylene promotes abscission, auxin inhibits abscission by maintaining the insensitivity of AL cells to ethylene (Nakano & Ito, 2013). Treatment with auxin on leaf blades or flowers can block the initiation of abscission, while organ abscission can be activated by auxin deficiency treatments through providing inhibitors of polar auxin transport, such as N-1 naphthyphthalamic acid (NPA), suggesting that auxin plays an important role in controlling abscission (Meir et al., 2006; Meir et al., 2010).

Other than ethylene and auxin, the involvement of other phytohormones, such as ABA, JA, cytokinins and gibberellins, in the regulation of abscission, has been controversial (Nakano & Ito, 2013). ABA was believed to directly play a role in abscission, but now is mostly associated with its ability to trigger tissue senescence. The effect of ABA on abscission appears to need interaction with auxin or ethylene, so this means that ABA only indirectly affects abscission (Patterson, 2001; Roberts et al., 2002).

A large number of genes are associated with the control of cell separation at the AL. When physiological changes, such as senescence or maturation, occur in organs, abscission is induced through reduction of the auxin level and induction of ethylene signalling (Taylor & Whitelaw, 2001; Meir et al., 2010). Abscission is carried out by activation of cell wall degrading enzymes and remodelling proteins, including ß-1,4-glucanase (cellulase), polygalacturonase, xyloglucan endotransglucosylase/hydrolase, and expansin, resulting in breakdown of cell adhesion (Roberts et al., 2002; Tucker et al., 2007; Cai & Lashbrook, 2008; Lashbrook & Cai, 2008; Meir et al., 2010). These events are well conserved in different types of organs and in divergent plant species, including both monocots and dicots.

1.3 Genetic mechanism of the shattering trait in monocots

Reduction in seed shattering is a significant trait acquired during crop domestication. A number of genes responsible for the seed shattering trait have been identified and cloned from several monocot species.

In rice, two significant genes, *qSH1* and *SH4*, that are implicated in seed shattering, have been identified. Konishi et al. (2006) stated that *OsqSH1* (*QTL of seed shattering in*

chromosome 1) is the major shattering QTL in rice (*Oryza sativa*) and has the largest effect on the shattering trait, explaining 68.6% of the total phenotypic variation in the population. The reason why domesticated rice lost its seed shattering trait is that a mutation identified as a SNP (Single Nucleotide Polymorphism) occurred in the 5' end regulatory region of *OsqSH1* (Konishi et al., 2006). This changed the rice phenotype from shattering to non-shattering, by negatively affecting abscission layer formation. However, Konishi et al. (2006) did not find a distinct ORF (Open Reading Frame) in the SNP region, but found one ORF located 12 kb away from the SNP. The ORF is a rice orthologue of the arabidopsis *REPLUMLESS* (*RPL*) gene (Roeder et al., 2003; Dinneny & Yanofsky, 2005; Konishi et al., 2006). The *RPL* gene encodes a BEL1-type homeobox which is involved in the formation of the dehiscence zone alongside the valve in the arabidopsis fruit (silique) (Reiser et al., 1995; Ito et al., 2002). The authors also demonstrated that the full shattering function of *OsqSH1* needs both the ORF and the SNP region and that the identified SNP only stops the formation of the abscission layer by changing the expression level of *OsqSH1* (Konishi et al., 2006). A homologue of *OsqSH1* was identified on chromosome 5 in rice, and named *OsSH5* (Yoon et al., 2014). They noted that *OsSH5* is highly expressed in the abscission layer in the pedicels and can enhance the seed shattering trait, but it still required *OsqSH1*. *OsSH5* was shown to induce another two key genes involved in seed shattering, *OsSHAT1* and *OsSH4*, determined in the pedicel region of *SH5*-overexpressed lines (Yoon et al., 2014).

The wheat *TaqSH1* has been cloned and characterised from homoelogous genome B and this gene was transformed into arabidopsis to investigate its function (Zhang et al., 2013). The results indicated that *TaqSH1* affected both silique dehiscence and floral perianth separation, which means that *TaqSH1* probably plays a universal role in the development of cells with abscission zone nature.

Li et al. (2006) showed that *SH4*, a QTL on rice chromosome 4, explained 69% of the total phenotypic variance relating to the seed shattering trait within an F2 population derived from a cross between *O. sativa* ssp. *indica* and the wild annual species *O. nivara*. *SH4* encodes a transcription factor, a member of the trihelix family, and is probably involved in the rice abscission layer development. It is located in a 1.7-kb region of a gene on rice chromosome 4 and is a dominant shattering gene in the wild species. Li et al. (2006) showed that the non-shattering (*sh4*) phenotype is from a simple nucleotide change in the first exon, leading to the amino acid asparagine being substituted by lysine (K79N) in cultivated rice. Their results

showed that *SH4* plays a key role in the formation of the abscission layer in early stages of flower development. However, as the expression level of *SH4* increased until the late stage of seed maturation, *SH4* may also take part in the activation of the abscission process (Li et al., 2006). Based on crossing experiments with *qsh1* and *SH4*, Onishi et al. (2007) demonstrated that *qSH1* is genetically epistatic to *SH4*.

Lin et al. (2007) identified *Shattering1* (*SHA1*) in introgression lines derived from the *indica* rice cultivar and a perennial common wild rice. *SHA1* is a allele of *SH4* and shares 98% amino acid identity. *sha1* also harbours the same lysine at residue 79 as *sh4*, which Li et al. (2006) identified as crucial to changing the function of the gene. However, there was no significant difference in abscission layer development in either the introgression line (*SHA1*) or the domesticated rice cultivar Teqing (*sha1*), as the abscission layer was fully developed, indicating that *SHA1* did not play a role in abscission layer formation (Lin et al., 2007), but Li et al. (2006) indicated that *SH4* is involved in abscission layer development. Lin et al. (2007) identified two more differences between *sha1* and *SH4*: firstly, the residue valine (V) in *sha1* at the position 152 is changed to residue alanine (A) in *SH4*; and secondly, six nucleotides (TGGAAA) from position 158 to 163 in *sha1*, are not present in *SH4*, suggesting that both differences in *SH4* are probably responsible for its role in abscission layer formation. However, Lin et al. (2007) considers *SHA1* is responsible for the degradation of the cell wall in the abscission layer, rather than abscission layer development.

Another *SH4* mutation, *sh4-2*, was identified by Zhou et al. (2012). s*h4-2* is a frame shift mutation that is apparently a null allele with a much stronger non-shattering phenotype, thus confirming the role of *SH4* in the shattering pathway. Zhou et al. (2012) demonstrated that that *SH4* appears to play a role upstream of *SHAT1* and *qSH1* in early abscission layer formation.

Recently, a quantitative trait locus, *GL4*, an orthologue of *SH4*, was identified from African cultivated rice (*Oryza glaberrima*). One SNP was identified in *GL4*, which resulted in a premature stop codon, and caused both smaller seed size and loss of seed shattering during African rice domestication (Wu et al., 2017). The mutation of *gl4/Obsh4* displayed an incomplete abscission layer. Wu et al. (2017) indicated that *GL4/SH4* is a key domestication gene with pleiotropic effects controlling both grain size and shattering in rice.

Zhou et al. (2012) also identified the *SHATTERING ABORTION 1* (*OsSHAT1*) gene, which is essential for seed shattering in rice by impacting on the differentiation of the seed abscission layer. *OsSHAT1* is an AP2 (APETALA2)-type transcription factor, the likely orthologue of the *AtAP2* in arabidopsis. The only AP2 gene reported affecting seed shattering is the wheat *Q* gene which has an impact on a range of characters important for domestication, such as glume shape and glume tenacity (Doebley, 2006; Simons et al., 2006). This means that the two important cereal crops may have analogous mechanisms for the initiation of abscission layers (Zhou et al., 2012). Interestingly, the expression level of *OsSHAT1* increased with the development of the spikelet. Zhou et al. (2012) proposed that *OsSHAT1* may be positively regulated by its gene products, based on the possibility that rice shares a similar mechanism with *AtAP2*. *AtAP2* participates in a self-feedback loop through negatively regulating its inhibitor, miR172 (Yant et al., 2010).

In contrast to *qSH1*, *SH4* and *SHAT1,* which promote abscission layer formation, *SH-H*, located in the 34 kb region of the chromosome 7, acts as a repressor of the seed shattering process in rice (Ji et al., 2010). *SH-H* is the only recessive shattering gene identified in rice. *SH-H* encodes a nuclear carboxy-terminal domain (CTD) phosphatase-like 1 (OsCPL1) protein. One universal trait of the CTD phosphatases in metazoans is related to cell differentiation. This CTD phosphatase domain plays a role in the dephosphorylation of the CTD of the largest subunit of RNA polymerase II (RNAP II) in eukaryotes (Ji et al., 2010). Progression of RNAP II through the transcription cycle is regulated by the state of CTD phosphorylation and the specific site of phosphorylation within the consensus repeat. Thus, CTD phosphatases may regulate the transcription of target genes by dephosphorylating the CTD of RNAP II (Ji et al., 2010). Therefore, in rice, OsCPL1 may repress the expression of genes controlling abscission layer cell differentiation by dephosphorylating the CTD of RNAP II or other phosphatase substrates (Ji et al., 2010).

In contrast to rice, seed shattering in wild sorghum (*Sorghum virgatum*) appears to be controlled by a single gene, *SH1,* which encodes a YABBY transcription factor (Lin et al., 2012). Three mutations at the *sh1* locus are retained in domesticated sorghums, changing the seed shattering trait to non-shattering trait. The three haplotypes include variants at regulatory sites in the promoter and intronic regions of the *sh1* gene in the domesticated line Tx430; a 2.2-kb deletion of the *sh1* caused a truncated transcript that lacked exons 2 and 3 in the domesticated line Tx623; and a substitution from GT to GG at a splice site of the *sh1* gene in

the domesticated line SC265. The three distinct haplotypes are distributed among 121 non-shattering sorghum cultivars worldwide, suggesting that the three different *sh1* mutations were selected independently (Lin et al., 2012). These mutations changed the *sh1* encoded protein, stopping the formation of the abscission layer in the joint between the hull and pedicel, thereby leading to a loss of shattering (Lin et al., 2012).

Cereal crops such as rice, wheat, maize and sorghum have been domesticated over thousands of years. These crops were cultivated from different wild progenitors by different ancient human groups in different geographical zones, but they all underwent systematic and parallel morphological and physiological changes during the domestication process (Harlan, 1992). It is still debated whether these parallel changes during domestication share the same genetic mechanism. However, Lin et al. (2012) reported genomic regions corresponding to *SH1* in cereals were conserved including in rice (*Oryza sativa*), maize (*Zea mays*), foxtail millet (*Setaria italica*) and sorghum (*Sorghum bicolor*).

A rice QTL for shattering, corresponding to *SH1* in sorghum, was identified through comparative genome analysis. PCR amplification and Southern blotting comparing by a non-shattering mutant and the wild-type rice breeding line (NANJING 11), indicating that a >4 kb insertion occurred in *Ossh1*, resulting in low transcript level and a shattering-resistant phenotype (Fukuta & Yagi, 1998; Lin et al., 2012). *OsSH1* was added to a list of genes under strong artificial selection (He et al., 2011).

Lin et al. (2012) also conducted a QTL mapping experiment using a large maize-by-teosinte population formed by crossing a maize inbred, W22, and a teosinte (*Zea mays* ssp. *parviglumis*). A major QTL, explaining 23.1% of the phenotypic variation of shattering, was identified to a narrow interval (genetic distance of 1.4 cM) on chromosome 5 with high confidence, and another QTL explaining 3.5% of the phenotypic variation for shattering was identified on chromosome 1 using whole genome linkage scan. These two QTLs correspond to the regions containing *ZmSH1-1* and *ZmSH1-5.1*+*ZmSH1-5.2* (Lin et al., 2012).

These examples indicate that *SH1* genes for loss of seed shattering might have been under parallel selection during the domestication of sorghum, rice and maize (Lin et al., 2012), with the implication that *SH1* orthologues may exist in other monocot plants, including perennial ryegrass.

SpWRKY, a shattering gene, located on chromosome 1 in a wild sorghum (*Sorghum propinquum*), encodes a WRKY transcription factor (Tang et al., 2013). The *WRKY* is a supergene family in plants affecting physiological and developmental processes, and some *SpWRKY* orthologues have been found including in rice and arabidopsis. Compared with *SbWRKY* (orthologue in wild non-shattering *S. bicolor* races), a mutation in *SpWRKY* from ATG to ATT leads to a substitution from methionine (M) to isoleucine (I) in amino acid sequence. As a result, the SpWRKY protein is 44 amino acid residues longer than the SbWRKY. The part of the extended sequence adjusts the protein structure, which may impact its ability to bind DNA. Moreover, *SpWRKY* has been found to play a role in regulating cell wall biosynthesis genes. Low expression of *SpWRKY* could derepress the downstream cell wall biosynthesis genes and allow the deposition of lignin that differentiates the abscission layer in the seed-pedicel junction. Interestingly, the *SH1* gene is located within 300 kb of *SpWRKY* on chromosome 1. The closeness of the two genes indicates that they may interact with each other impacting on the shattering phenotype. These two genes may regulate some of the same downstream genes involved in lignin deposition and may play important roles in the shattering pathway (Tang et al., 2013).

The identification of two orthologues *SpWRKY* and *SbWRKY* in sorghum suggested *WRKY* played little or no role in the domestication of sorghum, but implied the evolution of the genetic architecture of shattering in the wild species (Figure 1.2) (Lin et al., 2012). Additionally, the identification of *SpWRKY* in wild sorghum demonstrated that seed shattering is under opposing selection in domesticated crops and their wild relative, which means that the non-shattering trait has tended to be selected by humans, but a greater propensity for shattering is favourable for seed dispersal in wild species (Lin et al., 2012).

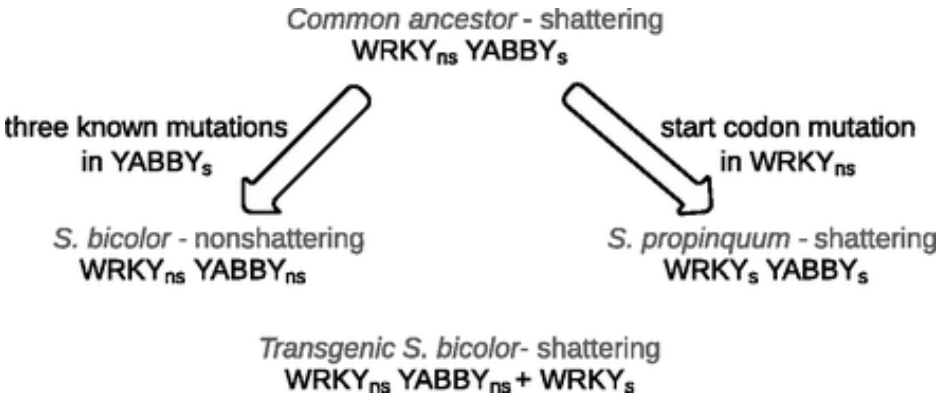


Figure 1.2 Inferred evolutionary history of two shattering loci in two sorghum species (Lin et al., 2012). Li et al. (2012) indicated that the ancestral sorghum that predated the divergence of *S. bicolor* and *S. propinquum* contained the shattering allele of YABBY but the non-shattering allele of WRKY. After the two sorghum species diverged, several non-shattering alleles of YABBY emerged in the *S. bicolor* population and were subsequently selected by humans, whereas the evolution of *SpWRKY* in *S. propinquum* potentially reinforced or further increased the level of shattering. $WRKY_{ns}$: *SbWRKY*; $WRKY_{s}$: *SbWRKY*; $YABBY_{ns}$: *SH1*; $YABBY_{s}$: *sh1*. ns: non-shattering; s: shattering.

Reproduced with permission from Haibao Tang, Hugo E. Cuevas, Sayan Das, Uzay U. Sezen, Chengbo Zhou, Hui Guo, Valorie H. Goff, Zhengxiang Ge, Thomas E. Clemente, and Andrew H. Paterson (2013) Seed shattering in a wild sorghum is conferred by a locus unrelated to domestication. PNAS 110:15824-15829: © 2013 National Academy of Sciences.

Ishii et al. (2013) reported that a mutation in a single locus (*SPR3*) resulted in a simple morphological change in panicle shape from open in wild rice (*Oryza rufipogon*) to closed in cultivated rice (*O. sativa*) and has a large impact on seed shattering and pollination. They demonstrated that plants with closed panicles retained mature seeds for about one day longer than those with open panicles. They detected the *SPR3* locus in a 9.3-kb genomic region, but no coding sequences were predicted in the region. The authors then identified that the rice liguleless gene (*OsLG1*) was located 10 kb away from the *SPR3* locus, encoding an SBP (*SQUAMOSA* promoter-binding protein) domain, that controls laminar joint and ligule development. They demonstrated that *OsLG1* is associated with the upstream 9.3 kb region of *SPR3*, and *SPR3* locus contains the regulatory sequence of *OsLG1* (Ishii et al., 2013). Some fixed polymorphic sites in the SPR3 locus were detected in *Oryza sativa* and wild *Oryza rufipogon* and transcription factor binding motifs were located in the fixed polymorphic sites (Ishii et al., 2013). Therefore, Ishii et al. (2013) speculated that one mutation may exist in one binding site, repressing the expression of *OsLG1* in the basal parts of the primary branches.

In summary, based on previous studies in rice, *SHAT1*, *SH4*, *qSH1* and *SH5* are essential genes involved in abscission layer differentiation in rice. However, knowledge about the regulation of these key genes is still limited. Zhou et al. (2012) proposed a genetic regulatory network (Figure 1.3), specifying AL development in rice. They suggested that *SH4* plays a role in the early stage of the AL formation. Based on the evidence that the earliest accumulation of *SH4* in the AL and inability to detect expression of *SHAT1* and *qSH1* in the *sh4* mutant in AL, Zhou et al. (2012) suggested that *SH4* acts largely upstream of *SHAT1* and *qSH1*. *SHAT1* appears to play a dual-functional role in the AL differentiation, including functioning downstream of *SH4* to activate *qSH1* expression and maintaining the expression of *SH4* in the AL. Zhou et al. (2012) indicated that *qSH1* affects the regulation of *SHAT1* and *SH4* expression in the AL, thus promoting the AL differentiation process. However, *qSH1* may not be the only genetic partner determining the expression of *SHAT1* and *SH4* in the AL, as some *qSH1*-defective rice sub-species appeared to show a reduced-shattering habit (Zhou et al., 2012).

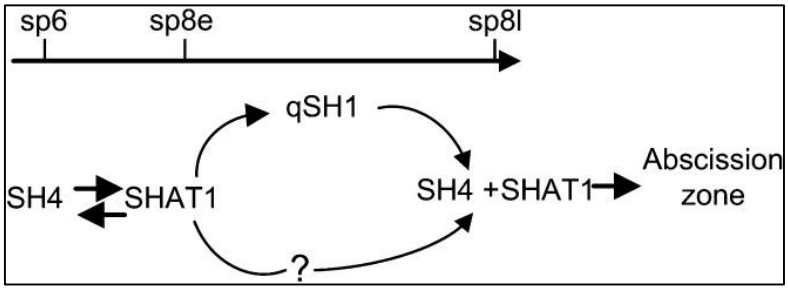


Figure 1.3 A Genetic Model of Regulatory Network Specifying abscission layer Development in Rice (Zhou et al., 2012). The long horizontal arrow represents spikelet developmental stages. sp8e, early stage sp8; sp8l, late stage sp8.

The most recent model for seed shattering in rice was presented by Yoon et al. (2017), as shown in Figure 1.4. Recently, Yoon et al. (2017) revealed that the KNOX protein, OSH15, encoded by *OSH15*, interacts with the BEL1 homeobox qSH1 and SH5. In addition, Yoon et al. (2014) indicated that both *qSH1* and *SH5* positively affect *SHAT1* and *SH4*, and *qSH1* and

SH5 have independent functions to maintain downstream genes. They suggested that the OSH15-qSH1 dimers induce AL differentiation by maintaining *SH4* and *SHAT1* and that OSH15-SH5 dimers inhibit the genes involved in lignin biosynthesis, which suppresses lignin accumulation in the AL (Yoon et al., 2017).

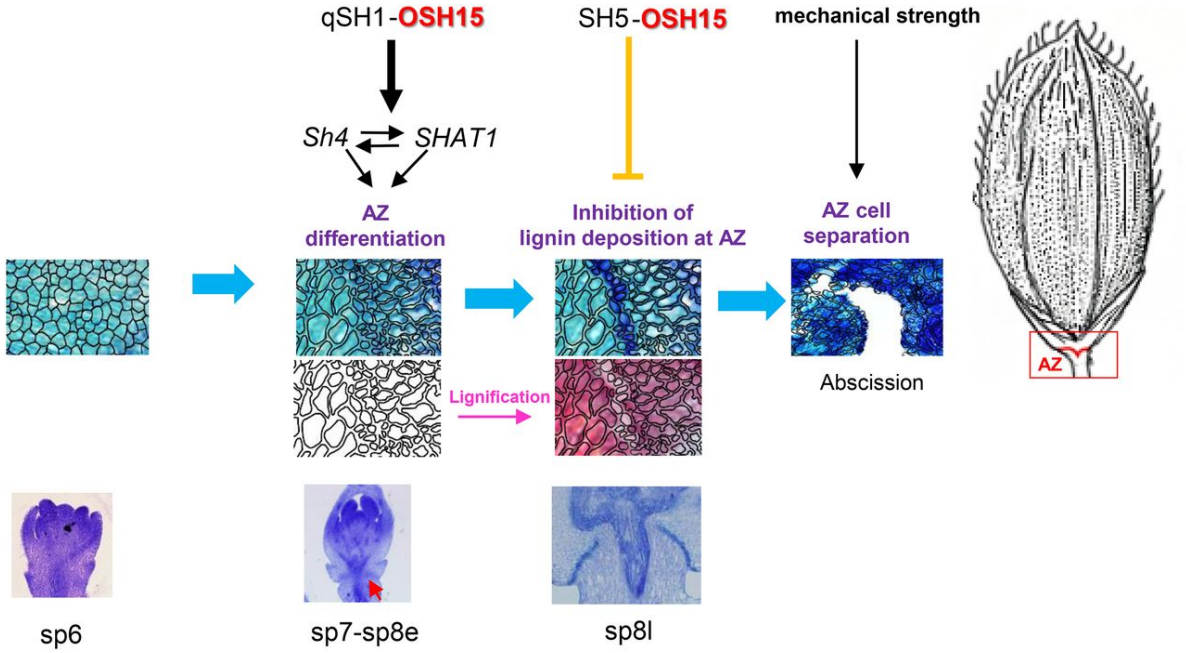


Figure 1.4 The most recent genetic model for the abscission process in rice (Yoon et al., 2017). The interaction between OSH15 and qSH1 promotes AZ development, upstream of SH4 and SHAT1. The OSH15-SH5 complex appears to suppress lignin accumulation. sp: spike development; sp8e, early stage sp8; sp8l, late stage sp8. AZ: abscission zone; has the same meaning as AL (abscission layer).

1.4 Cytokinins and plant productivity

Both seed size and seed number are the key factors contributing to seed production. Based on previous studies, the cytokinins could play a role in both seed traits in cereals (Ashikari et al., 2005a; Song et al., 2012; Zhang et al., 2012; and reviewed in Jameson & Song, 2016).

The cytokinins are a group of plant hormones that were first identified as promoting cell division and proliferation (Miller et al., 1955). The first natural cytokinin, zeatin, was isolated from the immature endosperm of *Zea mays* (Letham, 1963). Since then, the cytokinins have been implicated in the control of many aspects of plant growth and development, including shoot and root apical meristematic activity (Muller & Sheen, 2008; Shimizu-Sato et al., 2009), lateral shoot growth (Moubayidin et al., 2009), delay of leaf senescence (Guo & Gan, 2014), and seed development (Ashikari et al., 2005a; Zalewski et al., 2010; Jameson & Song, 2016).

Usually, zeatin is the most abundant naturally occurring free cytokinin: it has an isoprenoid moiety attached to an N(6)-substituted adenine derivative (Hothorn et al., 2011). The rate limiting step in cytokinin biosynthesis is the attachment of the isoprenoid side chain to ADT/ATP by isopentenyl transferase (IPT) (Sakakibara, 2006). The free base forms (tZ, cZ, DZ and iP) are considered to be the active forms (Lomin et al., 2015). Cytokinin metabolism occurs through modification of the purine moiety or by side chain modification or cleavage. The structural changes can be reversible or irreversible with complete loss of cytokinin activity. The details of reversible and irreversible conversions of cytokinin were reviewed by Spichal (2012), and an outline of cytokinin biosynthesis, interconversions and degradation in plants drawn by Spichal are shown in Figure 1.5.

Irreversible side chain cleavage by cytokinin oxidase/dehydrogenase (CKX) is a key method of regulating cytokinin content. To date, sequences of genes encoding CKX proteins have been identified in plants, including arabidopsis, maize, rice, brassica and wheat. The expression patterns of *IPT* and *CKX* gene family members have been identified in different plants, including in arabidopsis (Werner et al., 2003; Miyawaki et al., 2004), maize (Brugière et al., 2003; Brugière et al., 2008), brassica (O'Keefe et al., 2011; Liu et al., 2013; Song et al., 2015) and wheat (Song et al., 2012). In arabidopsis, *AtIPT* and *AtCKX* gene family members present overlapping expression patterns (Hirose et al., 2008). Increased expression of *CKX* genes and CKX activity has been shown to positively correlate with increases in cytokinin content in maize kernel (Brugière et al., 2003). It has been suggested that CKX activity plays a controlling role in determining seed number and seed size and consequently, impacts seed yield (Brugière et al., 2003; Bartrina et al., 2011; Li et al., 2013).

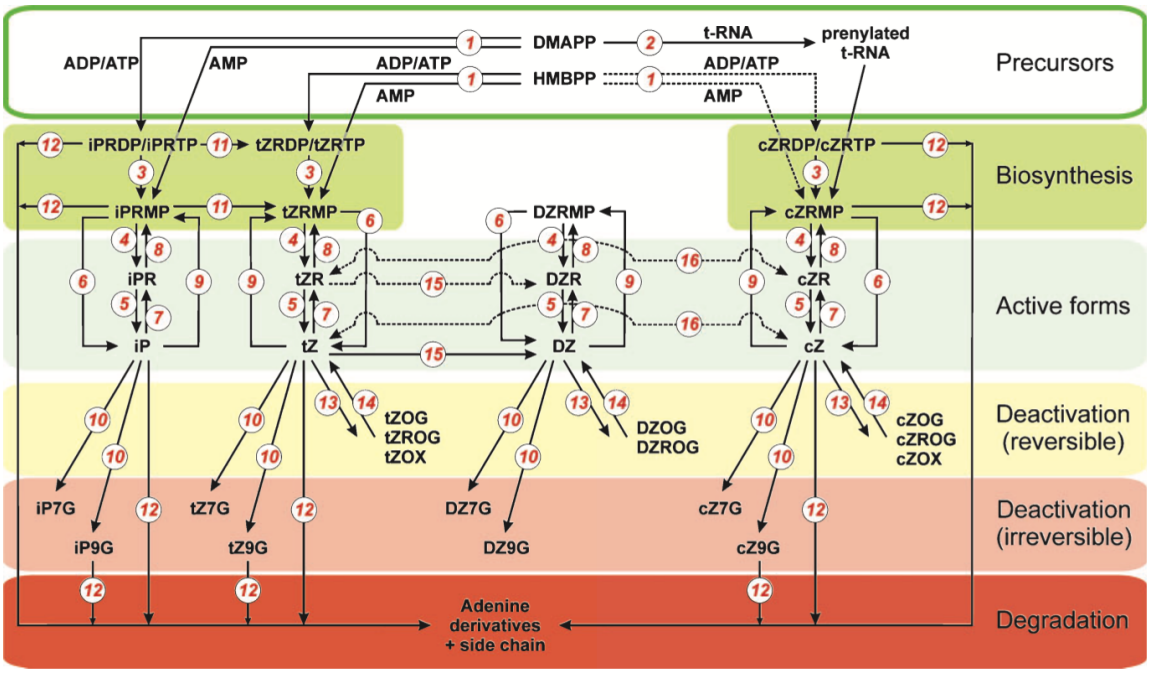


Figure 1.5 Scheme of CK biosynthesis, interconversions and degradation in plants. Enzymes involved in these interconversions are indicated by red numbers. Dashed lines show pathways that have not yet been sufficiently proven. (1) adenylate isopentenyltransferase (EC 2.5.1.27); (2) tRNA-specific isopentenyltransferase (EC 2.5.1.8); (3) phosphatase (EC 3.1.3.1); (4) 5′-ribonucleotide phosphohydrolase (EC 3.1.3.5); (5) adenosine nucleosidase (EC 3.2.2.7); (6) CK phosphoribohydrolase 'Lonely guy'; (7) purine nucleoside phosphorylase (EC 2.4.2.1); (8) adenosine kinase (EC 2.7.1.20); (9) adenine phosphoribosyltransferase (EC 2.4.2.7); (10) *N*-glucosyl transferase (EC 2.4.1.118); (11) cytochrome P450 mono-oxygenase; (12) cytokinin dehydrogenase (EC 1.5.99.12); (13) zeatin-O-glucosyltransferase, either *trans*-zeatin-specific (EC 2.4.1.203) or *cis*-zeatin specific (EC 2.4.1.215), utilising xylose instead of glucose (EC 2.4.2.40); (14) β-glucosidase (EC 3.2.1.21); (15) zeatin reductase (EC 1.3.1.69); (16) zeatin isomerase. DMAPP, dimethylallylpyrophosphate; HMBPP, 4-hydroxy-3-methyl-2-(*E*)-butenyl diphosphate; iPRDP, N^6-isopentenyladenosine-5′-diphosphate; iPRTP, N^6-isopentenyladenosine-5′-triphosphate; iPRMP, N^6-isopentenyladenosine-5′-monophosphate; iPR, N^6-isopentenyladenosine; iP7G, N^6-isopentenyladenosine-7-glucoside; iP9G, N^6-isopentenyladenosine-9-glucoside, and the equivalents for tZ, DZ and cZ; tZOG, *trans*-zeatin-O-glucoside; tZROG, *trans*-zeatin-O-glucoside riboside and the equivalents for DZ and cZ; tZOX, *trans*-zeatin-*O*-xyloside; cZOX, *cis*-zeatin-O-xyloside. Figure obtained from Spichal (2012).

Note: cytokinin ribosides (cZR, tZR, DZR and iPR) have no significant hormonal activity and are not active forms (Lomin et al., 2015).

Seed size and seed number are significant components of yield in cereals and both factors play important roles in seed sink strength (Smidansky et al., 2002). Seed number is directly related to the reproductive shoot apical meristem (SAM) controlling inflorescence meristem

development (Bartrina et al., 2011; Galli & Gallavotti, 2016). The activity of the SAM is regulated by many factors, including transcription factors and plant hormones (Veit, 2009). A study with transgenic tobacco where *CKX1* was overexpressed showed a reduced size of SAM, indicating that the cytokinins have a crucial role in SAM development and low levels of cytokinin and a reduction of cytokinin signalling can negatively affect activity of the SAM (Werner et al., 2008). Conversely, the *ckx3 ckx5* mutant of arabidopsis formed more siliques, 40% more than that of wild type, indicating that cytokinin plays a positive role in reproductive development (Bartrina et al., 2011).

Ashikari et al. (2005a) identified a QTL, *Gn1,* which explained 44% of the difference in grain number between Habataki (*indica* rice variety) and Koshihikari (*japonica* variety). The region of *Gn1a* carries a mutation in *OsCKX2*. Ashikari et al. (2005a) showed that one of mutations in *OsCKX2*, an 11 bp deletion in the coding region of *OsCKX2*, led to a premature stop codon and was a null mutation for *Osckx2*. Plants carrying this mutation produced over 400 grains in the main panicle in the field, 25% greater than other varieties, suggesting that a reduction or loss of function of OsCKX2 enhanced seed number (Ashikari et al., 2005a). This null mutation reduced the expression level of *OsCKX2* and elevated the cytokinin levels in the reproductive SAM, leading to increased activity of SAM and a consequent increase of grain number (Ashikari et al., 2005a).

The growth of the seed is related to the growth of the endosperm, and the endosperm mainly contributes to the size of the mature seed in monocots (Sundaresan, 2005). Endogenous cytokinin levels have been shown to have a positive correlation with cell division in seed development in maize, wheat and rice (Morris et al., 1993). For example the endogenous cytokinin in wheat sharply increased during the phase of rapid endosperm nuclear and cell division in the developing grain (Jameson et al., 1982; see also references in Jameson & Song, 2016). An orthologue of *OsCKX2* in wheat, *TaCKX6-D1* (phylogenetic alignment is to *TaCKX2* in Song et al. (2012)) was determined to be significantly related to grain weight (Zhang et al., 2012). Song et al. (2012) demonstrated that *TaCKX1*, an orthologue of *HvCKX1* in barley, *ZmCKX1* in maize and *OsCKX1* in rice, expressed strongly and specifically during early seed development stages. A study with barley lines where the expression of *HvCKX1* was repressed by RNAi technology, showed higher grain yield due to increases in seed number per plant and seed size (1000 grain weight) with increased seedling root mass (Zalewski et al., 2010). Song et al. (2012) suggested that regulating cytokinin

content during the seed development stages, such as down-regulating the expression level of *TaCKX1* to elevate the cytokinin level, may enhance seed number and seed size.

1.5 Reverse genetics and TILLING

Reverse genetics is a method used to discover the function of a gene by analysing the phenotypic effects of an induced gene disruption. In the last two decades, reverse genetics has developed rapidly, mainly because next generation sequencing (NGS) is more accessible, resulting in more knowledge about the genetic resources in various crops (Chen et al., 2014; Manzanares et al., 2016). It is possible that TILLING (Targeting Induced Local Lesions IN Genome) is the most adaptable reverse genetics method. TILLING combines traditional chemical mutagenesis (usually ethyl methane sulfonate, EMS) with a high-throughput mutation screening technique, to identify individuals in a population that contain allelic variation in a target gene (Till et al., 2003). An alternative to chemical mutagenesis, EcoTILLING focuses on the use of a wild or diverse genetic populations (L. Comai et al., 2004). Common polymorphisms without sample pooling can be discovered and catalogued by using EcoTILLING. For EcoTILLING, an equal amount of DNA from a reference sample is added to each assay well (Till et al., 2006).

TILLING has been applied to crop improvement because useful mutations can be used in traditional breeding programs, it is non-transgenic, and the novel trait is heritable (Dong et al., 2009a; Uauy et al., 2009; Kumar et al., 2013). Additionally, heterozygous recessive mutations, usually missed in phenotyping, can be identified by TILLING (Wilde et al., 2012). The TILLING approach has been applied to many crop species including rice (Wu et al., 2005; Till et al., 2007), wheat (Slade et al., 2005; Dong et al., 2009a; Chen et al., 2012; Acevedo-Garcia et al., 2017), maize (Till et al., 2004a), barley (Caldwell et al., 2004; Lababidi et al., 2009), sorghum (Xin et al., 2008; Blomstedt et al., 2012), potato (Muth et al., 2008), peanut (Knoll et al., 2011) and tomato (Minoia et al., 2010). Moreover, with developments in biotechnology, the mutation screening methods used in TILLING have been improved (Dong et al., 2009a; Parry et al., 2009; Minoia et al., 2010; Acevedo-Garcia et al., 2017).

In addition to TILLING, reverse genetics approaches include transgenic techniques, such as T-DNA insertion, RNA interference (RNAi) and genome editing (Chen et al., 2014; Manzanares et al., 2016). However, these techniques have drawbacks. Transgenic techniques

require time-consuming vector construction and plant transformation and the success rate of each case is variable (Fu et al., 2007). The current methods for plant transformation remain inefficient for many crops and forages species (Altpeter et al., 2016). The main challenges are long tissue culture periods and the small amount of DNA delivered by *Agrobacterium*-mediated gene transfer. In addition, there is no simplified transformation protocol: the transformation protocol may even need to be optimised for different cultivars from the same species (Bajaj et al., 2006), such as perennial ryegrass. Additionally, long periods of time in a de-differentiated state in tissue culture gives rise to somaclonal variants that may likely confound assigning phenotype to the induced mutations.

As there was no TILLING population of perennial ryegrass available, and as the public opinion in New Zealand is not favourable toward transgenic crop development, the EcoTILLING approach was used in this thesis, targeting seed shattering and seed size.

1.6 Aims and objectives

Seed shattering in perennial ryegrass is causing significant economic losses for the New Zealand seed industry. This project aimed to identify the key genes involved in the seed shattering process, and then to screen mutations of the target genes in a perennial ryegrass EcoTILLING population. As *CKX1* has been identified in rice and wheat to have a role in seed yield, potentially increasing seed size and number, it was decided to target mutations in *LpCKX1* as well.

The thesis is divided into a number of experimental Chapters:

- **Chapter 2**

The purpose of the experiments outlined in Chapter 2 was to develop a rapid and cost effective protocol for plant genomic DNA isolation. Isolation of high-quality genomic DNA is necessary, because EcoTILLING requires multiple DNA extractions, but the application of the CTAB method and the use of commercial silica column kits is restrained by time and cost when large numbers of DNA samples are required. A new protocol was established, combining these two methods, to reduce cost and save time.

- **Chapter 3**

Over the past decade, genes negatively controlling the shattering process have been isolated and their function shown in several crop species, including *SHAT1* and *SH4* in rice and the

SH1 gene in sorghum. For this thesis, it was hypothesised that perennial ryegrass has a similar mechanism to other Poaceae in terms of the control of seed shattering and that ryegrass shattering genes could be orthologous to genes detected in rice and sorghum. The aim of the experiment outlined in Chapter 3 was to identify the key candidate genes putatively involved in the seed shattering trait in perennial ryegrass.

- **Chapter 4**

The aim of the experiment outlined in Chapter 4 was to optimise techniques for mutation screening in a perennial ryegrass EcoTILLING population. In the TILLING approach, mutations are usually detected with gel electrophoresis. A new method combining HRM analysis to detect CEL I digestion products was developed to reduce the time taken for mutation screening. The new method to detect CEL I digestion products was tested to screen mutations in *LpCKX1*.

- **Chapter 5**

The aim of the experiment outlined in Chapter 5 was to screen for mutations in the target genes by using HRM analysis using an optimised homemade HRM master mix.

2 Chapter 2: DNA extraction method

This chapter has been published: Fu Z-Y., Song J-C. and Jameson PE. (2017) *A rapid and cost effective protocol for plant genomic DNA isolation using regenerated silica columns in combination with CTAB extraction.* Journal of Integrative Agriculture 16(8): 1682-1688. http://dx.doi.org/10.1016/S2095-3119(16)61534-4. (Journal Articles)

2.1 Introduction

Isolation of high quality genomic DNA is necessary for many molecular biology applications. For plant genomic DNA isolation the popular method using CTAB (hexadecyltrimethylammonium bromide) is cost effective with high DNA yield and acceptable DNA quality, but this protocol is time consuming (Doyle & Doyle, 1987; Allen et al., 2006). On the other hand, high quality DNA can be isolated rapidly using commercial DNA extraction kits with easy protocols and then the used columns are discarded (Deavours & Dixon, 2005; Tesniere et al., 2006). However, the application of either the CTAB method or the use of commercial kits can be limited both in terms of time and expense when large numbers of DNA samples are required. As silica matrices are extraordinarily stable over extended time periods under mild acid conditions, it has been suggested that silica columns may be reused after acid treatment to remove any DNA carried over on the binding matrix (Siddappa et al., 2007). The authors reported on the successful use of regenerated columns to purify plasmids. However, there are few attempts to isolate plant genomic DNA using regenerated silica columns. Lemke et al. (2011) reported on the regeneration and reuse of Qiagen 'DNA Easy 96 column plates' and columns, but these were regenerated using a commercial regeneration kit.

To detect seed shattering mutants or cytokinin oxidase/dehydrogenase mutants in perennial ryegrass requires analysing leaf material from multiple single seed lines from numerous cultivars using EcoTILLING (Luca Comai et al., 2004; Till et al., 2006; Song et al., 2015). In this chapter, a rapid and cost-effective plant genomic DNA isolation protocol is reported that meets the requirements for preparation of large numbers of genomic DNA samples for such applications as genotyping and large-scale mutation screening. In this method, regeneration of commercial silica columns (Qiagen DNeasy Plant Mini Kit, Cat. No. 69104) was assessed. To test their reliability, silica columns regenerated over 10 cycles were used for plant genomic DNA isolation in combination with a modified CTAB method. Retention capacity

and quality of eluted DNA were assessed, as well as using PCR to test for cross contamination when isolating genomic DNA from two different species.

2.2 Materials and Methods

2.2.1 Column regeneration

New Qiagen DNA extraction columns (Qiagen DNeasy Plant Mini Kit, Cat. No. 69104) were used to isolate genomic DNA from *Lolium perenne* (perennial ryegrass) following the manufacturer's protocol. The used columns were then washed thoroughly with water to remove any cell debris, and soaked in either 0.5 M or 1 M HCl for either 1, 4, 24 or 48 h. The columns were then rinsed thoroughly with sterile distilled water 3-5 times, and 500 µl of buffer QBT were added. The QBT buffer was prepared in-house based on the recipe supplied in the Qiagen Plasmid Purification Handbook and comprised 750 mM NaCl, 50 mM MOPS (3-(N-Morpholino)propanesulfonic acid) (pH 7.0), 15% (v/v) isopropanol and 0.15% (v/v) Triton X-100. The columns were spun at 13,000*g* for 1 min. The columns were ready for a fresh application and then used for up to 11 rounds of regeneration following the above steps. Two sets of controls were used. These were new columns used to isolate perennial ryegrass genomic DNA, but the columns were either not soaked or soaked in H_2O for 1 h, 4 h, 24 h or 48 h. Elution buffer was applied to the columns and the eluate collected and treated as if it contained DNA. Eluted DNA was assessed using a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), and by gel electrophoresis on 1% (w/v) agarose gel prepared in 35 ml 1x TAE buffer (40 mM Tris-acetate, 1mM EDTA) and 2 µl SYBR safe.

2.2.2 DNA extraction procedure

Lysis buffer (2% CTAB buffer with 2% PVP40 (w/v); 2% β-mercaptoethanol (v/v) added just before use.) was pre-heated to 85°C. Up to 100 mg fresh weight (FW) of plant material was ground to a fine powder under liquid nitrogen using a pre-chilled mortar and pestle. Alternatively, when an Omni International Bead Ruptor 24 (Omni International, Kennesaw, Georgia, USA) was used, 100 mg tissue and 3 ceramic beads (2.8 mm Ceramic Bead Media 19-646-3; Omni International) were placed into a 2 ml freestanding microtube and stored in liquid nitrogen. Samples were disrupted using the bead ruptor with speed set at 3.9 m/s for 20 s. The powdered tissue was then scraped into a dry Axygen 1.7 ml microtube and 400 µl

preheated lysis buffer was added, and the mixture incubated at 85°C for 10 min. [Note: this is a modification of the standard CTAB procedure]. After 5 min incubation, the tubes were gently shaken for 5 s to disperse the material, and the incubation continued. The sample was spun at 20,000*g* for 5 min at room temperature; 300 µl of the supernatant was transferred into a new 1.5 ml tube and mixed with 500 µl of binding buffer (2 M guanidine hydrochloride, 75 % (v/v) ethanol) and then 700 µl of this mixture was transferred onto new or regenerated columns. The columns were spun at 13,000*g* for 1 min at room temperature. The flow-through was discarded. 700 µl of washing buffer I (10 mM NaCl, 10 mM Tris-HCl pH 6.5, 80 % (v/v) ethanol) was applied to each column, and the column centrifuged again at 13,000*g* for 1 min. This step was repeated once. The columns were then washed with 700 µl of washing buffer II (96% v/v ethanol) and spun at 13,000*g* for 1 min. The flow-through was discarded and the column centrifuged again at 13,000*g* for 2 min. 50 µl of elution buffer (comprising 10 mM Tris-HCl pH 8.5, with RNase A (Qiagen Cat. No.145012547) was added to a final concentration of 10 µg/ml), was preheated to 65°C, and applied to the column. The columns were spun at 13,000*g* for 1 min to elute DNA. This step was repeated once using the elution buffer from the collection tube to obtain a greater DNA concentration. Alternatively, addition of a further 50 µl of elution buffer could be added and the column centrifuged again which would potentially increase yield, but reduce the DNA concentration. The eluted DNA was incubated at 37°C for 1 h to remove any remnants of RNA, and the DNA products stored at -20 °C.

2.2.3 Yield and quality determination

In order to determine the yield and quality of DNA obtained from regenerated columns, a single, pre-treated sample from 1 g FW perennial ryegrass leaves was used for the whole experiment. This was done so that identical amounts of initial DNA could be applied to each test column. The leaves were ground to a fine powder under liquid nitrogen, the powder scraped into a dry 15 ml tube and 4 ml lysis buffer, preheated to 85°C, was added. The samples were incubated at 85°C and, because of the large volume, incubation time was extended to 18 min to lyse cells adequately. The solution was then spun at 20,000*g* for 3 min at room temperature. The supernatant was transferred to a new tube. Following this, 300 µl of the mixture was transferred into new Axygen 1.7 ml microtubes and binding buffer (500 µl) was added; 700 µl of the mixture was then applied to 12 silica columns. The DNA was eluted as described above. The 12 columns included a new column and columns that had been

regenerated between 1- and 11-times (Regenerations 1 to 11). The experiment was repeated three times.

2.2.4 Assessing DNA quality using restriction enzymes

The concentration of perennial ryegrass genomic DNA extracted using new and regenerated columns was adjusted to 200 ng/µl. The DNA was cut with *BamH* I and *Hind* III (Thermo scientific, Cat. No. ER0051 and Cat. No. ER0501, respectively). Each assay contained 2000 ng of DNA, four units of enzyme and 1x digestion buffer (with restriction enzyme), and was incubated for 1 h at 37°C. The reaction products were resolved by electrophoresis on 1% (w/v) agarose gels with TAE buffer.

2.2.5 Quality check using amplification of genomic DNA by RT-qPCR

DNA collected from new, 1-, 5- and 10-times regenerated columns was tested. PCR primers for the elongation factor (EF) gene (sequence shown in Table 2.1) were used for RT-qPCR that was carried out on a Rotor-Gene Q (Qiagen Helden, Germany). A homemade 2x SYBR master mix was used for the RT-qPCR (Song et al., 2012). Each reaction contained 100 ng of DNA. The cycler program was: 95°C for 10 min followed by 40 cycles of 95°C for 10 s, 58°C for 15 s and 72°C for 20 s. After amplification, a melting program was followed, with the temperature increasing from 72 to 95 at 1°C every 5 s. Ct values and the melt curve analysis were generated by the Rotor-Gene Q Software (Qiagen).

Table 2.1 Primers sequences

Name	Sequence	Amplicon size (bp)
LpEF-F	CACCCTGGTCAGATCGGCAAC	238
LpEF-R	CACCAACAGCAACAGTCTGCCT	
PsGAP-F	TCTCTTCGGTCAGAAGCCAGTTAC	256
PsGAP-R	GCAGCTAGCATTGGAAATAATGTCAAAC	
LpGAP-F	AGGAGGTTGCYGTSTTTGGCTG	237
LpGAP-R	TAGCRTTRGAGACAATGKYGATGTCAGA	

2.2.6 Testing for DNA cross contamination

To test whether there was carry-over of DNA in the regenerated columns, columns that had been used to isolate genomic DNA from perennial ryegrass and regenerated up to 5 and 10

times were used. Genomic DNA from *Pisum sativum* (pea) was applied to either a new column or to columns regenerated for five and 10 rounds (Re5 and Re10). The column purified DNA was then tested by PCR, using GAPDH (glyceraldehyde 3-phosphate dehydrogenase) primers specific for pea (Dhandapani et al., 2016; Jameson et al., 2016) and perennial ryegrass (Roche et al., 2016) to determine if there was any perennial ryegrass DNA from the regenerated columns contaminating the pea DNA. Primers sequences are listed in Table 2.1.

The PCR program was: 94°C for 10 min followed by 35 cycles of 94°C for 20 s, 58°C for 30 s, 72°C for 30 s and a final amplification step at 72°C for 6 min. The PCR products were run on a 1% (w/v) agarose gel with TAE buffer.

2.3 Results and discussion

When multiple DNA extractions are needed the cost of silica columns can become prohibitive. If the silica columns can be reused, the cost can be substantially reduced. However, there is a significant risk if the DNA from previous rounds is not completely removed and is carried over in the silica column to contaminate a subsequent extraction, or if the capacity of the column to retain DNA is substantially reduced.

2.3.1 Modification of the CTAB method

Based on the traditional CTAB method, 2% CTAB buffer was used to lyse the plant cells but at 85°C, instead of at 60°C (Doyle & Doyle, 1987) and successfully reduced the incubation time from 40 to 10 min. The integrity of DNA from samples incubated at 85°C for 10 minutes was similar to DNA from samples incubated at either 65°C or 75°C for 20, 30 or 40 min. The DNA product from the 85°C incubation was not degraded within 10 min and the yield and quality were still excellent (Figure 2.1). However, a longer incubation at 85°C (e.g., for 20 min) led to degradation of the DNA (Figure 2.1). In some cases, for example, when the DNA quality is more important than speed, an intermediate incubation temperature (60-65°C) should be tested.

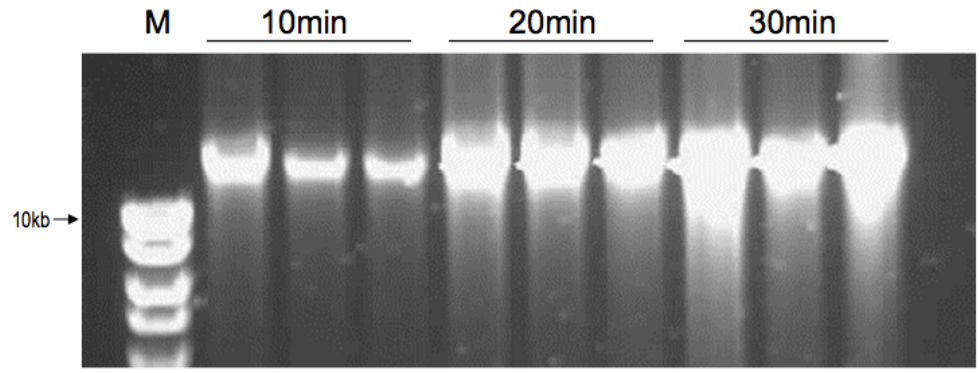


Figure 2.1 Modification of CTAB method to reduce lysis incubation time. Ground samples, in three replicates, were incubated in lysis buffer pre- heated to 85°C for 10, 20, and 30 min. The DNA was purified through Qiagen DNeasy columns and the products applied to 1% (w/v) agarose gel with TAE buffer. M, molecular weight marker (HyperLadder 1 kb, Bioline, UK).

2.3.2 Column regeneration

To determine the optimal method for cleaning the columns, HCl was used as first proposed by Siddappa et al. (2007). As shown in Figure 2.2, following application of perennial ryegrass DNA to the columns, all DNA removed when the columns were soaked in HCl, even in 0.5 M HCl for as little as 1 h. In Siddappa et al. (2007), plasmid DNA cannot be visualized on the 1% agarose gel after 4 h incubation in 1 M HCl and DNA binding capacity of columns was not reduced by prolonging soaking time in 1 M HCl for as long as 30 days. It is still possible that a trace of DNA could be in the columns after 1 h, but not visible on the gels, so it is proposed that columns are soaked in 1 M HCl overnight (16 h), with the columns then stored in 0.1 M HCl. The columns must then be rinsed with dH_2O and buffer QBT before use.

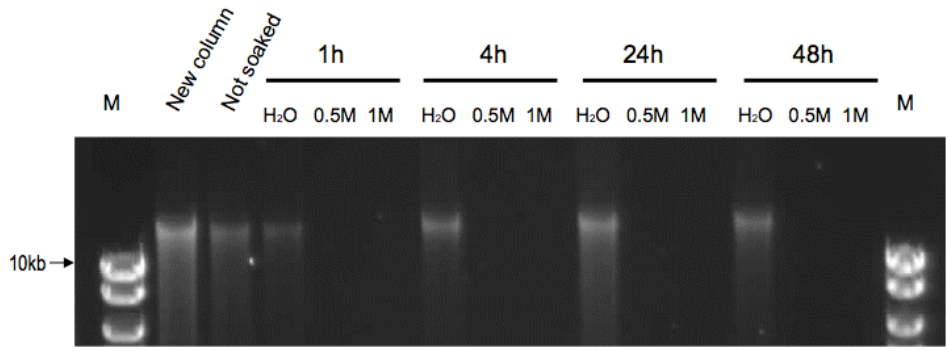


Figure 2.2 Optimization of the regeneration protocol. After perennial ryegrass DNA was applied and eluted, the used columns were soaked in 0.5 M HCl, 1 M HCl or H_2O for 1, 4, 24, and 48 h. The soaked columns were washed with distilled water twice and QBT buffer once, and then eluted with 50 µl elution buffer. The controls included a used column without soaking and a new column to which DNA had been applied. All elution products were run on a 1% (w/v) agarose gel with TAE buffer. M, molecular weight marker (HyperLadder 1 kb, Bioline, UK).

2.3.3 Yield and quality of DNA were not compromised when regenerated columns were used

Using the above regeneration protocol, the binding capacity of the columns was tested to determine if the binding capacity declined with increasing numbers of regeneration. There was some impact on binding but, even after 11 regenerations, the product concentration was still about 300 ng/µl (Figure 2.3A), which is sufficient for most experiments, including PCR and RT-qPCR. Furthermore, the integrity of the DNA was maintained. Figure 2.3B shows that the product was not degraded. In addition, when the purified DNA was assessed spectrophotometrically, acceptable values of between 2 and 2.2 for ratios of 260/280 nm and 260/230 nm respectively, were obtained (Appendix 2.1).

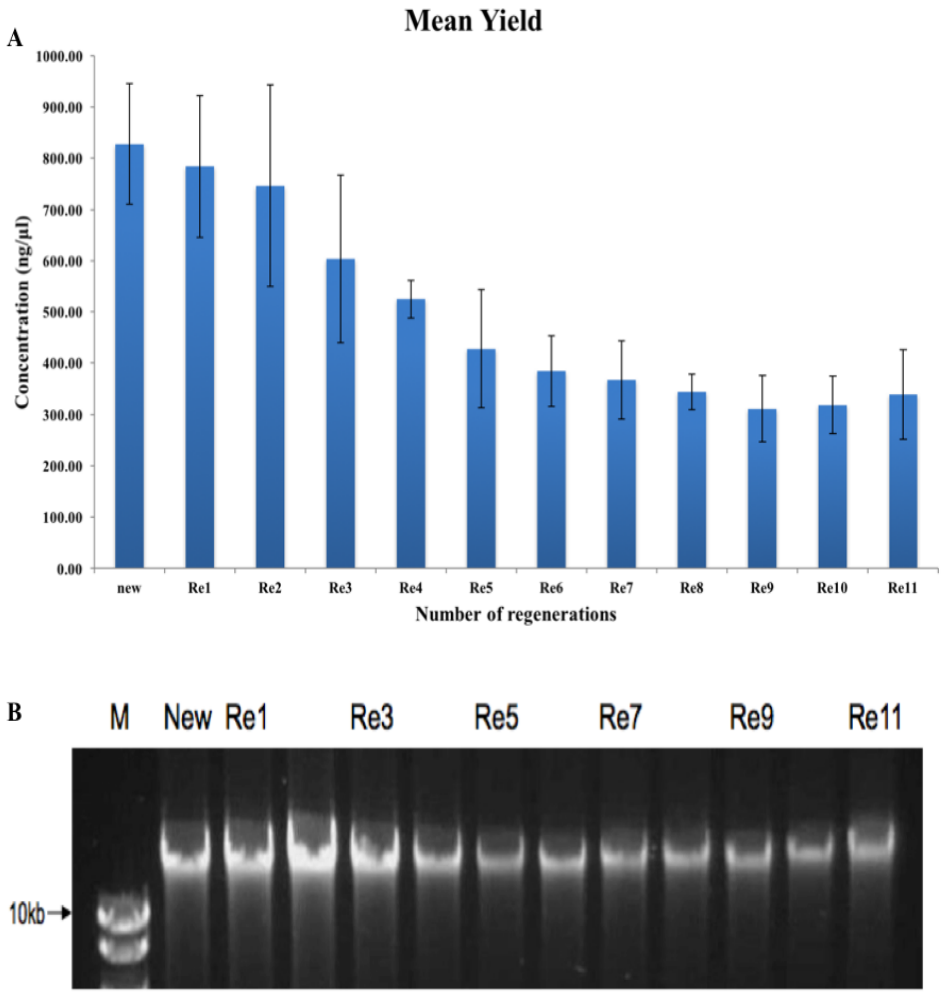


Figure 2.3 Comparison of yield and integrity of DNA obtained from a new column and columns regenerated up to 11 times (Re1–11). Perennial ryegrass DNA was added to the pre-prepared, regenerated columns and new columns. DNA products were detected by NanoDropND-1000 (A) and gel electrophoresis on 1% (w/v) agarose with TAE buffer (B). M, molecular weight marker (HyperLadder 1 kb, Bioline, UK). The DNA yield is presented as the mean of three replicates and the error bars are ±SD.

A further test of the DNA quality included assessing the effectiveness of cutting by restriction enzymes. Restriction digestion is a commonly used molecular technique, which utilises restriction endonucleases. Two frequently used restriction enzymes, *BamH* I and *Hind* III, were used to test the quality of genomic DNA isolated by the protocol using new and regenerated columns. As shown in Figure 2.4, the perennial ryegrass genomic DNA, from new, Re1, Re5 and Re10 columns, showed good integrity prior to enzyme digestion, but after 1 h of digestion, the DNA appeared as a smear on the gel, indicating that the DNA was

completely cut by both *BamH* I and *Hind* III. This indicates that high quality DNA was produced even after the columns had been regenerated multiple times.

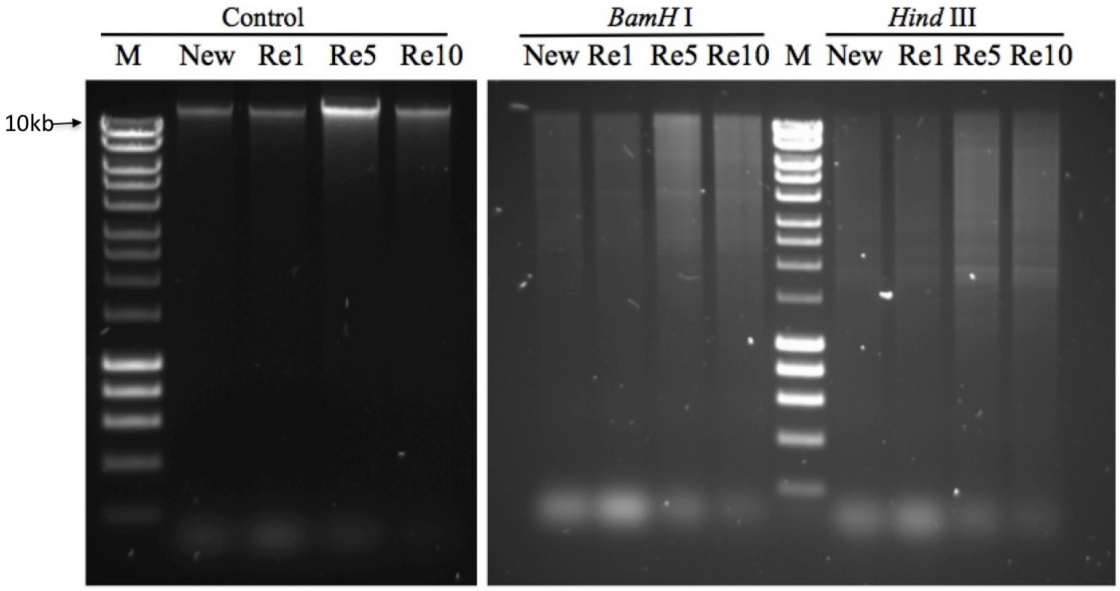


Figure 2.4 Quality assessment using restriction enzyme assay of DNA eluted from a new column and columns regenerated up to 11 times (Re1–11). Perennial ryegrass DNA was applied to a new column, and columns regenerated for 1 (Re1), 5 (Re5) and 10 (Re10) times. The eluted samples were cut by *BamH* I and *Hin*d III, respectively. The DNA products without digestion and the digestion products were run on a 1% (w/v) agarose gel with TAE buffer. M, molecular weight marker (HyperLadder 1 kb, Bioline, UK).

As isolated genomic DNA is frequently used for cloning, the quality of the DNA was also checked by amplification of genomic DNA using RT-qPCR. The purity of the genomic DNA is critical for PCR analysis, so the DNA from different regenerated columns was used for RT-qPCR. RT-qPCR can only be used reliably when the amplification is without error, so the quality of the template is critical to success. The fluorescence curves and the melt curve analysis (Figure 2.5) show that the amplifications were performed successfully and the products were of high specificity. Based on these results, the DNA from different columns (new, Re1, Re5, Re10) is of high quality.

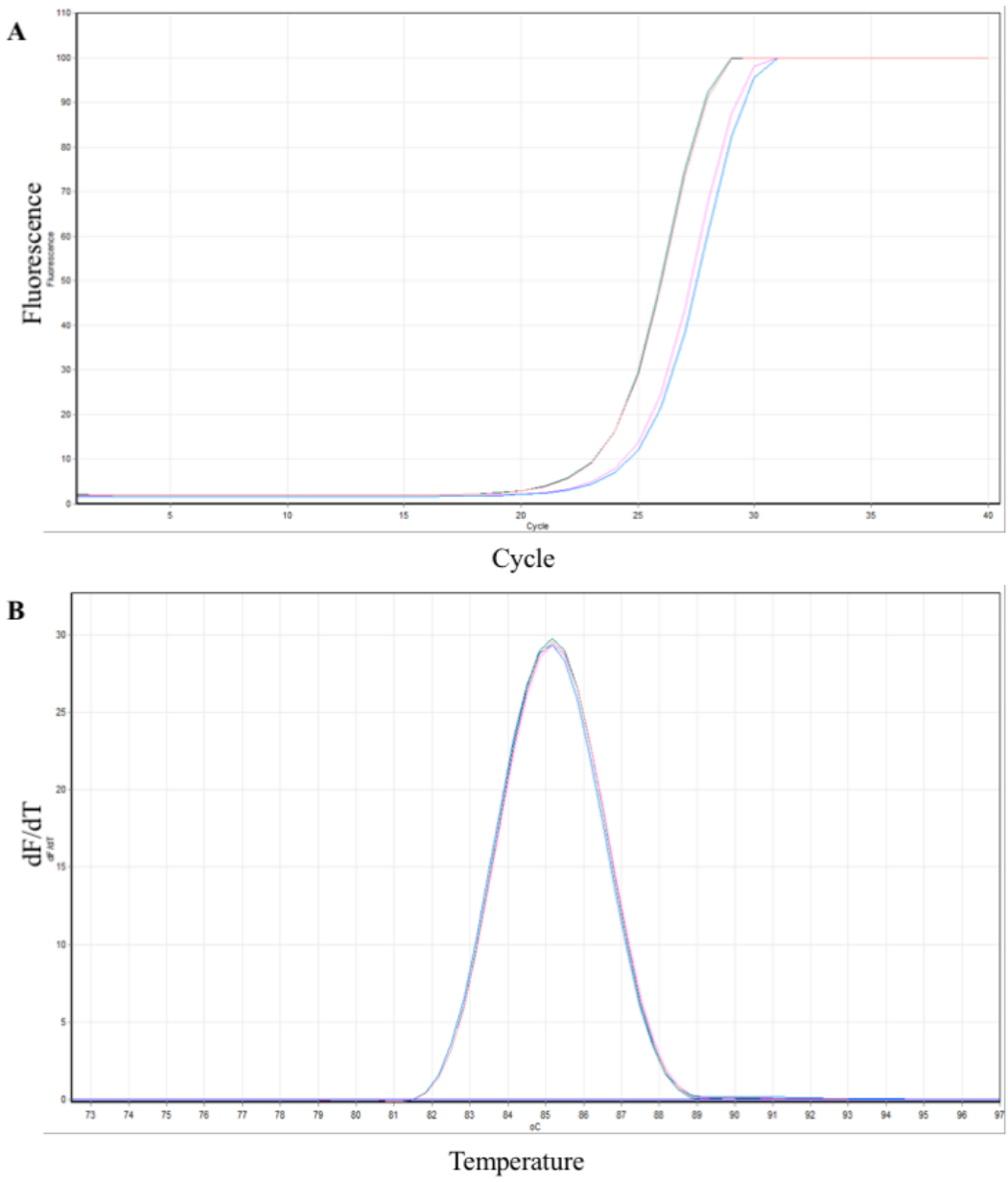


Figure 2.5 Amplification of genomic DNA with RT-qPCR. The DNA are products from New, Re1, Re5, and R10 using the procedure described in Fig. 4. The fluorescence curves (A) and the melt curve (B) analysis were generated by the Rotor-Gene Q software 2.1.0.9.

2.3.4 Sample-to-sample cross contamination did not occur

When HCl was used to clean columns, the result was acceptable visually (Figure 2.2). However, it was still possible that there may have been traces of DNA carried over in the regenerated columns, leading to cross contamination. To test this, columns that had previously been used to isolate DNA from perennial ryegrass were used to isolate genomic DNA from pea. As Figure 2.6 shows, 5th and 10th regeneration columns were used to isolate

pea DNA having previously been used to purify perennial ryegrass DNA. The DNA products eluted from the columns were amplified by using pea and perennial ryegrass gene-specific primers. Positive controls were run for both perennial ryegrass and pea DNA using new columns. The PCR products from the regenerated columns (previously used for perennial ryegrass DNA) were the same as from the new column to which only pea extract had been applied: no PCR product was detected using the perennial ryegrass primers, indicating that there was no perennial ryegrass contamination of the pea genomic DNA.

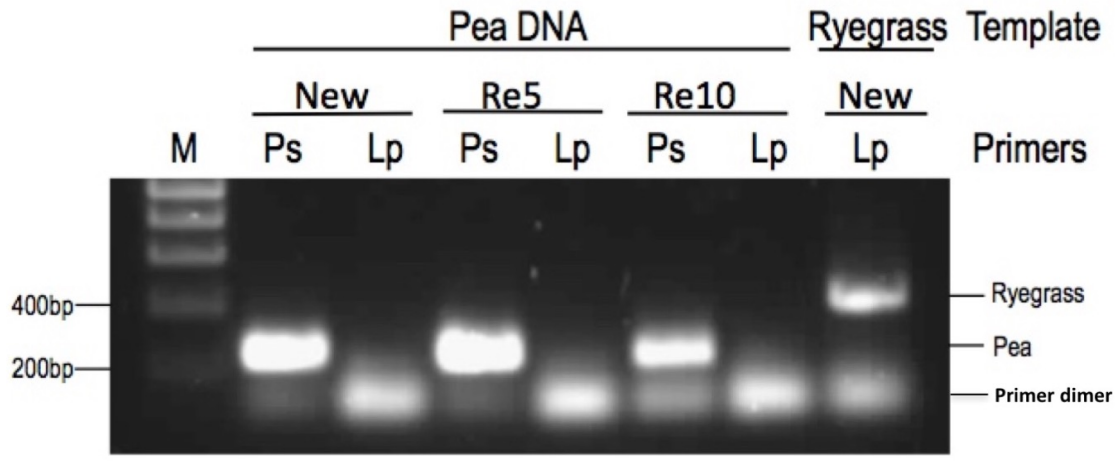


Figure 2.6 Comparison of PCR products after elution of pea genomic DNA from new and regenerated columns. The columns were initially used for perennial ryegrass DNA extraction and were regenerated 5 (Re5) and 10 (Re10) times. After that, new, Re5, and Re10 columns were used to extract pea DNA. Pea DNA was amplified by standard PCR with pea primers and perennial ryegrass primers. A control is perennial ryegrass DNA with perennial ryegrass primers. The PCR products were checked on 1% agarose gel with TAE buffer. M, molecular weight marker (HyperLadder 1 kb, Bioline, UK). Ps, pea primers; Lp, perennial ryegrass primers.

2.4 Conclusions

A rapid and cost effective protocol for isolating plant genomic DNA is described. This protocol not only combines the traditional CTAB method with commercial silica columns, but also improves the CTAB method to reduce incubation time, and re-uses regenerated silica columns up to 11 times. The reliability of the regenerated columns was tested for yield, quality and purity. High yield and good quality genomic DNA was produced by a low-cost

and facile method. In addition, this method used individual silica columns, providing greater flexibility and convenience than using 96-well column plates (Lemke et al., 2011). Based on these benefits, this protocol could meet the requirements for preparation of large numbers of genomic DNA samples for such applications as genotyping and large-scale mutation screening.

3 Chapter 3: Identification and expression of putative seed shattering genes in *Lolium perenne* L.

3.1 Introduction

The developmental stages of the inflorescence of perennial ryegrass is shown in Figure 3.1. Spike initiation commences with the appearance of spikelet buds in the axils of the leaf primordia, forming the typical 'double ridges' (Figure 3.1 A). While the spike is developing, leaf primodia at the base of the apex are still expanding to form leaves. The spike is surrounded by leaves until heading. The formation of florets as secondary buds on the spikelets is shown in Figure 3.1 B (Cooper, 1951). Each spike contains 20 to 30 spikelets attached to the rachis and spikelets constist of 5 to 13 normally developed florets. The florets are arranged on the rachilla in the same plane as that of the spikelets on the rachis, as shown in Figure 3.1 C and D (Elgersma et al., 1988).

The abscission layer joins organs and the main body of the plant together. In perennial ryegrass, Elgersma et al. (1988) suggested that the abscission layer is across the rachilla, just below each floret (Figure 3.1 D). In addition, Elgersma et al. (1988) indicated that the ryegrass abscission layer was formed before heading, and the mature seed started to disperse about 3.5 weeks after anthesis.

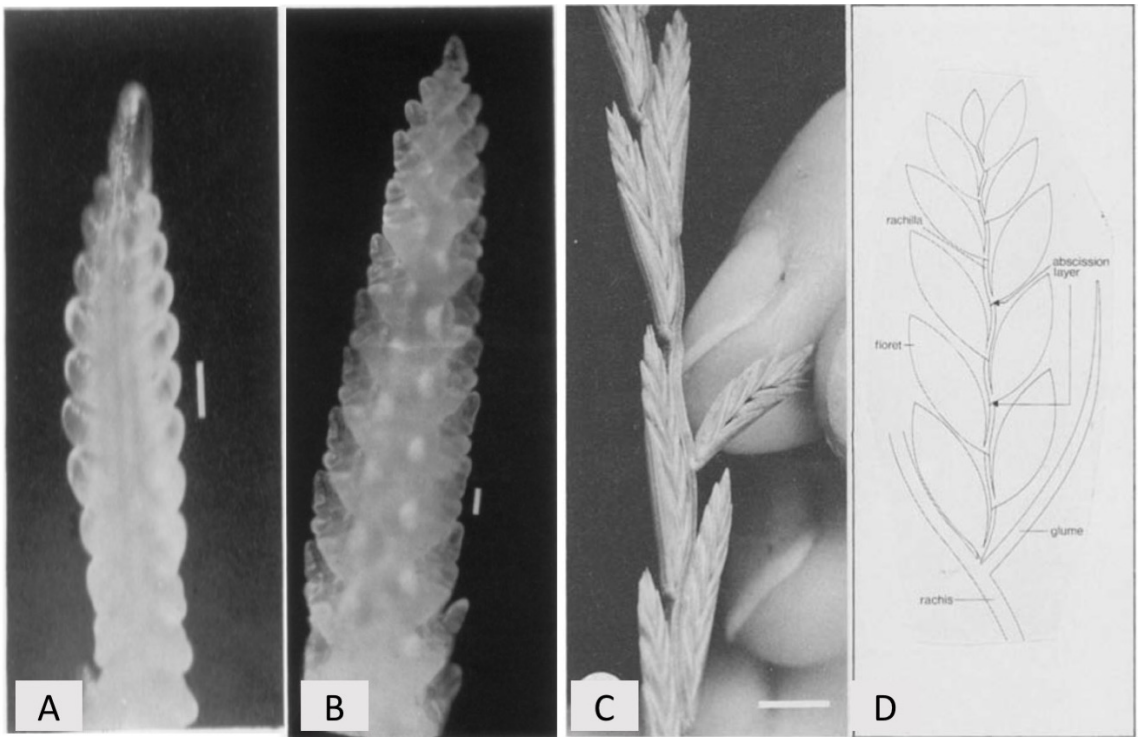


Figure 3.1 Spike and spikelet of ryegrass. A: Shoot apex, showing first appearance of spikelet buds (Cooper, 1951). B: young inflorescence, showing development of floret buds on the spikelets. Scale bar, 0.1 mm (Cooper, 1951). C: spikelets with florets. Scale bar, 2.5 cm (X 0.4) (Elgersma et al., 1988). D: a spikelet, showing the position of the abscission layers across the rachilla, just below each floret (Elgersma et al., 1988).

In the last decade, several genes involved in the seed shattering trait have been identified in different domesticated crops, including *SH4*, *SHAT1*, and *qSH1* in rice (Konishi et al., 2006; Li et al., 2006; Zhou et al., 2012), *Q* in wheat (Simons et al., 2006), and *SH1* in sorghum, rice and maize (Lin et al., 2012), as reviewed in Chapter 1. Seed shattering is a developmentally programmed abscission event, occurring in a specific tissue, the abscission layer (AL). Genes controlling seed shattering were generally dominant in the early progenitors of crops such as rice and sorghum (Konishi et al., 2006; Li et al., 2006; Lin et al., 2012). During domestication, non-shattering, homozygous recessive mutant alleles, replaced the dominant alleles as a consequence of unconscious selection over millennia (Konishi et al., 2006; Lin et

al., 2012). However, many forage plants, such as perennial ryegrass (*Lolium perenne.* L.), are much less domesticated than crops used for human food, having been actively bred for less than 100 years, so the seed shattering trait is still retained.

Functional analysis of the shattering genes indicates that they could share similar roles in the abscission process in different species, such as *SH1* in sorghum, rice and maize (Lin et al., 2012). Lin et al. (2012) demonstrated the syntenic relationships between *SH1* genes in sorghum, rice, maize, and foxtail. All these crops and ryegrass belong to the family Poaceae (Jones et al., 2002). Therefore, it is hypothesised that homologues of the shattering genes identified in other crops could play a similar role in perennial ryegrass.

To test the hypothesis that perennial ryegrass has a genetic mechanism similar to cereal crops for controlling the seed shattering trait, a comparative genomics approach was applied (Jones et al., 2002) through BLAST searching the NCBI database and a perennial ryegrass RNA-seq transcriptome database prepared by Professor Jiancheng Song (Roche et al., 2016; Guo et al., 2017), and the candidate seed shattering genes in perennial ryegrass were identified. The relative expression of the ryegrass candidate genes was then determined during different developmental stages of the reproductive tissues using reverse transcription quantitative PCR.

3.2 Material and methods

3.2.1 Plant material

Plant material was collected for three independent biological replicates for the gene expression study. In the first biological replicate, multiple seeds from *Lolium perenne* L. cv. RI009 were sown in several 2 L pots with fertilised soil at the University of Canterbury glasshouses (43° 31' 48" S, 172° 37' 13" E). For the first three months, the pots were placed outside the glasshouses for vernalisation and then removed to a glasshouse with 15-minute watering in the morning every two days. Temperatures in the glasshouse ranged from 15ºC minimum at night to 25ºC during the day. The material was collected from October to December, 2014, under increasing day length from 12.5 h to 15.5 h. For the spike collection by length, the early spike developmental stages were measured by gently feeling along the leaves covering the spike. If the length was in the target size range of 1-2 cm and 4-8 cm, the outside leaves were peeled off using a scalpel (Figure 3.2 A and B). When the spike was

longer than 10 cm and in the process of heading, the whole spike was collected. The spikelets were tagged on the day of first flowering and then collected at different days after anthesis. In addition, other vegetative tissues including roots and leaves of one-week old petri dish germinated seedlings, flag leaves and nodes, were also collected. Flag leaves and nodes were collected at about the heading stage. The samples were labelled and collected randomly to make sure samples from each stage were from different individual ryegrass plants. Samples were immediately plunged into liquid nitrogen.

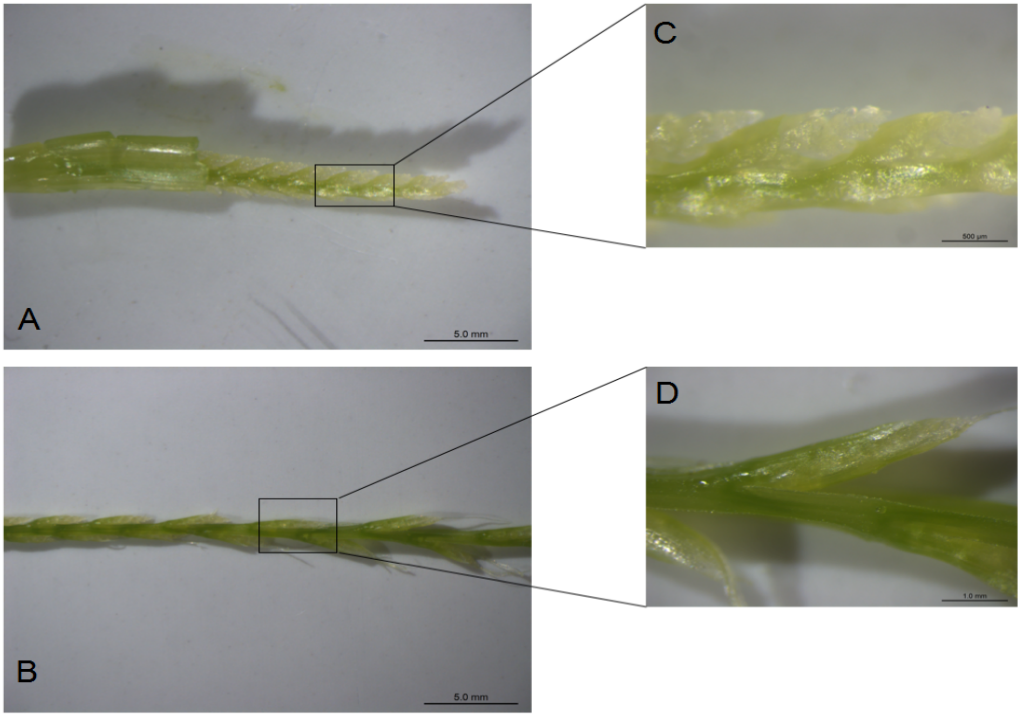


Figure 3.2 The early developmental stages of spikes. Two early stages of spike formation (A: spike 1-2 cm long; B: spike 4-8 cm long). These two stages were detected by finger touching, and then the flag leaf was peeled off and the whole spike collected. C, D: enlarged images of the boxed area in the left-hand panels, showing development of floret buds on the spikelets. Refer also to Tables 3.1 and 3.3.

For the second biological replicate, plants of cv. Grasslands Nui were grown in a field located west of Christchurch (43°34'00.5"S, 172°26'45.9"E) in 2014. Whole spikelets at different days after anthesis were collected from different individual ryegrass plants, and all materials for this biological replicate was initially collected by Professor Jiancheng Song. No early

spike and vegetative tissues were collected for this replicate. The harvested plant material was immediately placed in liquid nitrogen and stored at -80°C.

For the third biological replicate, plants of cv. Grasslands Nui were grown at Yantai University (37° 28' 28" N, 121° 27' 27" E), China. The seeds were sown on the ground in a glasshouse on February 2015, and were transplanted to outside of glasshouse when about one month old for vernalisation. Plants were watered every morning. The material was collected from May to June 2015, as the day length increased from 13.5 h to 14.5 h. The floral stages and vegetative tissues collected for the third biological replicate were the same as for the first biological replicate, and the samples at each developmental stage were collected from different individuals. The harvested plant material was immediately placed in liquid nitrogen and stored at -80°C.

For each of the three biological replicates, the dissection of the abscission layer was slightly different. For spikes between 1-2 cm long, the whole spike was used for RNA extraction. Whole spikelets were collected from the middle third of the spikes that were 4-8 cm long. For spikes longer than 10 cm and all samples after this stage, spikelets from the middle third of the spikes were dissected to collect the pedicel junction between the flower/grain and rachilla, the part which contained the abscission layer. The abscission layer dissection process was performed under liquid nitrogen. The upper part of each floret was cut and discarded, so the bottom of each floret and pedicel junction were collected, providing samples enriched in the abscission layer. More details are listed in Tables 3.1-3.3.

Table 3.1 Plant material collected for the first biological replicate for RNA extraction. Material at each stage was collected from different individuals.

Tissue	Developmental stage	AL collection
Spike	Spike length 1-2 cm	Whole spike
	Spike length 4-8 cm	Whole spikelet from middle third of spikes
Spikelet	Spike length >10 cm	AL [a]
	0 daa	
	2 daa	
	5 daa	
	10 daa	
	15 daa	
	20 daa	
Vegetative tissue	Seedling leaves	N/A
	Flag leaves	
	Node	
	Seedling root	

[a] The tops of each floret on the whole spikelet were cut off, so the pedicel junction and the rachilla of the whole spikelet were collected for RNA extraction. daa: days after anthesis.

Table 3.2 Plant material collected for the second biological replicate for RNA extraction. Material at each stage was collected from different individuals.

Tissue	Developmental stage	AL collection
Spikelet	-1 daa	AL [a]
	1 daa	
	4 daa	
	7 daa	
	14 daa	
	18 daa	
	21 daa	

[a] Two basal florets of each spikelet were collected and the tops of the florets were cut off to collect part of pedicel and rachilla. daa: days after anthesis.

Table 3.3 Plant material collected for the third biological replicate for RNA extraction. Material at each stage was collected from different individuals.

Tissue part	Developmental stage	AL collection
Spike	Spike length 1-2 cm	Whole spike
	Spike length 4-8 cm	Whole spikelet from middle third of spikes
Spikelet	Spike length >10 cm	AL [a]
	-2 daa	
	0 daa	
	4 daa	
	6 daa	
	10 daa	
	14 daa	
	16 daa	
	18 daa	
Vegetative tissue	Seedling leaves	N/A
	Flag leaves	
	Node	
	Seedling root	

[a] Two basal florets of each spikelet were collected and the tops of the florets were cut off to collect part of pedicel and rachilla. daa: days after anthesis.

3.2.2 RNA isolation and cDNA synthesis

All samples containing the AL were ground in 1.5 ml centrifuge tubes with plastic mini pestles in the presence of liquid nitrogen. The samples from vegetative tissues were ground using mortar and pestles with liquid nitrogen. Total RNA was extracted from up to 100 mg of frozen samples using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and DNase I set (Qiagen, Hilden, Germany) to avoid genomic DNA contamination following the manufacturer's instructions and immediately stored at -20°C. Total RNA from flag leaves was isolated by using Ambion TRIzol (Invitrogen, USA) and then purified by using a Qiagen clean-up kit (Qiagen, Hilden, Germany) following the user guide. The quality and integrity of isolated RNA were assessed on 1% (w/v) agarose gel (Invitrogen, USA) with 2 μl of Sybrsafe DNA gel stain (Invitrogen, USA). The concentration was assessed by NanoDrop-1000 spectrophotometer.

For the first and second biological replicates, approximately 1 μg of total RNA was converted to complementary DNA (cDNA) with 50 U Expand Reverse Transcriptase (Roche,

Mannheim, Germany) by following two steps. Firstly, for primer annealing, a 10 µl reaction contained 50 pmoles of oligo (dT) primers, 100 pmoles of random hexamer (pdN6) primers, 0.5 µl of RNAsecure (Life Technologies, USA) and 1 µg of extracted RNA was incubated at 65°C for 10 min and was then transferred immediately to ice. Secondly, for reverse transcription, 10 µl of the master mix containing 4 µl of 5x RT buffer, 1 µl of 20 mM dNTPs (Bioline, UK), 2 µl of DTT with the enzyme kit, 2 µl of H_2O, and 1 µl of Expand Reverse Transcriptase (50U/µl) was added. The reaction was then incubated at 25°C for 10 min, and then 42°C for 60 min, and finally the reaction was inactivated at 70°C for 15 min. For the third biological replicate, the reverse transcription reaction was carried out with cDNA synthesis kit (Clontech, Japan) following the manual. The cDNA products were diluted 10-fold with Milli-Q water and stored at -20°C.

3.2.3 Candidate seed shattering genes in perennial ryegrass

Candidate target gene sequences from different monocots were collected from the NCBI database and published articles. The collected DNA sequences of each target gene were aligned using the Clustal X software (Version 1.83) to verify their identity, and were used as query sequences to BLASTN search the perennial ryegrass transcriptome database[1] to determine the homologues in perennial ryegrass. This initial search was conducted by Professor Jiancheng Song. The homologues from perennial ryegrass were named *LpSH1*, *LpLG1*, *LpSH4*, *LpqSH1a*, *LpqSH1b*, *LpSHAT1a*, *LpSHAT1b*, *LpWRKY* and *LpQ*. The Neighbour-Joining (NJ) Phylogenetic trees were constructed using Clustal X software. The NJ phylogenetic tree was visualised with MEGA 7 software. The tree was rooted with an outgroup sequence from Ginkgo *AP2*. The resource of homologues of candidate genes are listed in the Table 3.4 including GenBank accession numbers and journal articles. All homologues of candidate shattering genes were renamed in this thesis to for convenient reading.

[1] The RNA-seq transcriptome database was constructed by Professor Jiancheng Song, containing 169 862 assembled sequence contigs of 595 bp in average length generated using an Illumina HiSeq2000 genome analyser at Macrogen, Korea.

Table 3.4 The sources of homologous genes described in this thesis

Candidate gene	Names used in this thesis	Species	Accession No./journal	Function
SH1	ZmSH1	Zea mays	NM_001148293	putative YABBY domain transcription factor family Protein
	BdSH1	Brachypodium distachyon	XM_003561722	protein YABBY 2-like
	PeSH1	Phyllostachys edulis	FP098112.1	not annotated
	HvSH1	Hordeum vulgare	AK358221.1	not annotated
	TaSH1	Triticum aestivum	AK332816.1	not annotated
	OsSH1	Oryza sativa	EU846982.1	YABBY mRNA
	SvSH1	Sorghum virgatum	Lin et al. (2012)	
	SiSH1	Setaria italica	Lin et al. (2012)	
	OsSH1	Oryza sativa	Lin et al. (2012)	
	ZmSH1-5.1+5.2	Zea mays	Lin et al. (2012)	
	ZmSH1-1	Zea mays	Lin et al. (2012)	
LG1	OsLG1	Oryza sativa	AB776991	LG1
	HvLG1	Hordeum vulgare	AM117950	liguleless-like protein
	ZmLG1	Zea mays	NM_001112073	LG1
	SbLG1	Sorghum bicolor	XM_002447206	protein LIGULELESS 1
	BdLG1	Brachypodium distachyon	XM_003560343	squamosa promoter-binding-like protein 10
SH4	OsSH4	Oryza sativa	JN697614	shattering 4
	EcSH4	Echinochloa crus-pavonis	AB455329	sh4 homologue
	ZmSH4	Zea mays	BT069839	not annotated
	SbSH4	Sorghum bicolor	XM_002448695	not annotated
	HvSH4	Hordeum vulgare	AK376279	not annotated
qSH1	OsqSH1	Oryza sativa	AB071330	qSH-1
	ZmqSH1	Zea mays	EU973659	not annotated
	SbqSH1	Sorghum bicolor	XM_002458717	not annotated
	BdqSH1	Brachypodium distachyon	XM_003564561	BEL1-like homeodomain protein 9
	HvqSH1	Hordeum vulgare	AK367206	not annotated
	TaqSH1	Triticum aestivum	AB546647	BEL1-type homeodomain protein
	LpqSH1-E1691	Lolium perenne	GR521691	not annotated
SH5	OsSH5	Oryza sativa	Yoon et al. (2014)	
Q	TaQ	Triticum aestivum	AY645945	AP2-like transcription factor; Q protein
SHAT1	LpSHAT1-E4971	Lolium perenne	GR514971	not annotated
	LpSHAT1-E4720	Lolium perenne	GR514720	not annotated
	OrSHAT1	Oryza rufipogon	FO82280	not annotated

	TaSHAT1	Triticum aestivum	AK331198	no annotated
	SbSHAT1	Sorghum bicolor	XM_002447175	no annotated
	PeSHAT1	Phyllostachys edulis	FP093799	no annotated
	HvSHAT1	Hordeum vulgare	AK371953	no annotated
WRKY	BdWRKY	Brachypodium distachyon	XM_003559570	probable WRKY transcription factor 71-like
	ZmWRKY	Zea mays	NM_001151284	putative WRKY DNA-binding domain superfamily protein
	ObWRKY	Oryza brachyantha	XM_006651590	probable WRKY transcription factor 2-like
	Hv WRKY	Hordeum vulgare	JQ806389	WRKY transcript factor 48
	OsWRKY	Oryza sativa	BK005063	WRKY transcription factor 60
	PeWRKY	Phyllostachys edulis	FP100870	not annotated
	SiWRKY	Setaria italica	XM_004982196	not annotated

3.2.4 Primer design and optimisation for gene expression study

All primers for the gene expression study, including reference genes and the candidate seed shattering genes, were designed using Primer Premier 6.2 software. Ideally primers targeting mRNA should span an exon boundary, and the primer length should be between 20-30 bases. Additionally, the GC content should be between 40-60%. Generally, two forward (F) primers and two reverse (R) primers were designed for each target gene, and one pair for each gene was then selected to perform quantitative PCR with the synthesised cDNA. All primers were synthesised by Macrogen, South Korea and the primer sequences are listed in Table 3.4.

Table 3.5 Primer sequences for gene expression study

Name	Sequences [a]
LpSH1F1	AGAACATATGCTACACATGCGACCAG
LpSH1F2	AAGAGGCAGCGGGTTCCTTCAG
LpSH1R1	GATGGCCTCGTCCAGCTTCTTG
LpSH1R2	GAGGAATATTCGTCTCAGTAGAGACCTTG
LpLG1 F1	TGGTGAAGGAAATGCAGGTGGATC
LpLG1 F2	AGGGGAAAGCATCAATGCAGCAG
LpLG1 R1	CAGGCTGTTCTTGCTGCACTTG
LpLG1R2	GACCCAGCTGAAGGATGTTGCTG
LpSH4F1	ACTACCGCAAGGGCAACTGGAC
LpSH4F2	ACGAGACGCTCGTCCTCATCAC
LpSH4R1	GTCGTTGCACTGGTTCTGGCTG
LpSH4R2	CGCGCGCACCTTCTTGTAGTC
LpqSH1aF1	GCTCGAGGAGATCTGCGACGTG
LpqSH1aR1	GTACCTCTTGCAAACCTCTTCCATCATG
LpqSH1aR2	GCCTGGACCTGCTGGTAGTACTG
LpqSH1aF2	ATGATGGAAGAGGTTTGCAAGAGGTAC
LpqSH1aF3	GGCAGTACTACCAGCAGGTCCAG
LpqSH1aR3	GATGTCTTTGCTCATGCCTTCTTTGAC
LpqSH1bF1	TGCTCATCTCGCTCATGGAAGAAGTT
LpqSH1bF2	GCTACAAGCAATACTACCAGCAGCT
LpqSH1bR1	CCCGAAGTTCGCCATCTCTTCCT
LpqSH1bR2	GCCTCCCATGAGCCCGAAGTTC
LpZSHATF	CGAGGATTACGAGGACATCAAGCA
LpZSHATR	TCCCATCTGCCGCACTTGTG
LpSHAT1aF1	GGACATCAAGCAGATGGGCAATCTG
LpSHAT1aF2	CTGACCAAGGAGGAGTTCGTCCAC
LpSHAT1aR1	GGCTTCCTCCTCGGTATCGAACAG
LpSHAT1aR2	CGTCCTTGCCGTTGCACTTGATG
LpSHAT1bF1	CGACGCCGACATAAACTTCAACCTC
LpSHAT1bF2	CGAGGACGACATGAAGCAGATGAAG
LpSHAT1bR1	ACTTCGTTGTCGAATAGCCCAAGATATATG
LpSHAT1bR2	TTGCAGCCTCTACTTCGTTGTCGA
LpQF1	GGATYTGAAGCAGATGAGGAACTGGAC
LpQF2	AGGAGGAGTTCGTGCACATCCTC
LpQR1	TCGTAAGCTCTTGCAGCTTCAACTTC
LpQR2	CTTCCCTCCCRTTGAAGCGAATG
LpWRKYF1	AGAGGACGGGTACCAGTGGAAGAAG
LpWRKYF2	AGTGGAAGAAGTACGGCCAGAAGTTC
LpWRKYR1	GACGGTGAGCCGTGCTGGTG
LpWRKYR2	TACTGGGCGCTGAGCTCGTAC
LpGAPF1	AGGAGGTTGCYGTSTTTGGCTG
LpGAPR1	GTCAGASKTGTAYTCCTTCTCRTTG
LpGAPR2	TGKYGATGTCAGASKTGTACTCCTTC
LpGAPR3	TAGCRTTRGAGACAATGKYGATGTCAGA
LpEFF1	CACCCTGGTCAGATCGGCAAC
LpEFF2	CCCACATCGCCGTCAAGTTCTC
LpEFF3	GCCGTCAAGTTCTCTGAGATCCAG
LpEFR1	CACCAACAGCAACAGTCTGCCT

[a] Expected product size of for primer selected for expression study are shown in Table 3.10.

3.2.5 Reference gene identification

Reference genes are typically constitutively expressed genes that are required for maintenance of basal cellular function, so these genes are minimally regulated in different

tissues and developmental stages of an organism. The expression level of the reference genes was also used as quality controls of cDNA samples. All cDNA samples with C_t (cross threshold) value difference of five or less were considered as good in quality and were used for expression studies in this work. Two forward primers and two reverse primers were designed for each reference gene, *GAPDH* (*GAP*) and *Elongation Factor* (*EF*), and were tested for expression stability and efficiency against a mixed cDNA from different developmental stages of the spike. One primer pair for each reference gene was selected for the relevant gene expression study and primer sequences are listed in Table 3.5. After optimisation, the reference genes were used to normalise cDNA samples across the three biological replicates.

3.2.6 Target gene sequence verification

For sequence verification of each gene of interest, the PCR reaction was performed using BioTaq DNA polymerase (Bioline, UK) with the mixed cDNA from first biological replicate. The 15 µl of PCR master mix contained 1 Unit of Taq DNA polymerase (Bioline, UK), 1.5 µl Bioline 10x PCR buffer, 0.2 mM of each dNTP (Bioline, UK), 0.66 µM of each primer and about 1 µl of mixed cDNA as template. The reaction was subjected to 94°C for 5 min, followed by 30 cycles of 94°C for 10 s, 58°C annealing for 15 s, and 72°C extension for 20 s, and final 72°C extension for 10 min. The PCR products were then separated on 3% (w/v) agarose gels with TAE buffer, at 7 V/cm for 40 min. The bands of the expected size (Table 3.5) were purified by using UltraClean 15 DNA Purification Kit (Mo Bio Laboratories, USA) following the instruction manual. The purified PCR products were directly sequenced by Macrogen Inc, Korea. The sequencing results were only for sequence verification, so some SNPs in the sequence from different alleles were acceptable.

However, for sequencing the putative *LpSH4* and *LpLG1* genes, some INDELs probably exist in both of these sequences, so the sequencing results are unreadable and cannot be used for sequence verification. Therefore, the PCR products were cloned using the TA-cloning method with pMD 18-T Vector (TaKaRa, Japan). The experimental process is described in Section 3.2.8. The plasmids were subsequently sequenced.

3.2.7 Reverse transcription quantitative PCR (RT-qPCR)

The expression of the genes of interest was performed using reverse transcription quantitative PCR. The primer pairs were tested and the best pair for each gene was selected. A homemade

SYBR Green (Invitrogen, USA) master mix was used for RT-qPCR in a Rotor-Gene Q real-time PCR instrument (Qiagen, Hilden, Germany). The components of the homemade master mix are listed in Tables 3.6-3.8. The reaction was subjected to 94°C for 10 min, followed by 40 cycles of 94°C for 10 s, 58°C annealing for 15 s, and 72°C extension for 20 s, and finally the melt temperature range was set from 72 to 95°C with a ramp setting at 1°C/2s. There were at least three technical replicates for each primer pair. The relative expression of candidate seed shattering genes was corrected and calculated using $2^{-\Delta\Delta Ct}$ method as described in previous studies (Pfaffl, 2001; Song et al., 2008).

Table 3.6 The components of the 2x qPCR buffer

Component	Volume/Amount per 1 mL
Trehalose	227 mg
H_2O	340 µl
10x NH_4 Buffer [a]	200 µl
50 mM $MgCl_2$	160 µl
DMSO	80 µl
10% Triton-X 100	30 µl

[a] PCR buffer in the BioTaq DNA Polymerase reagent kit (Bioline, UK)

Table 3.7 The components of 2x homemade SYBR Green master mix

Component	Volume per 1 mL
2x qPCR buffer [a]	930 µl
20 mM dNTPs	40 µl
Bioline Taq	15 µl
100x SYBR Green [b]	15 µl

[a] 2x qPCR buffer recipe in Table 3.6

[b] 100 times diluted with DMSO from SYBR Green I Nucleic Acid Gel Stain (Invitrogen, USA)

Table 3.8 Reaction composition using 2x SYBR Green homemade master mix

Component	Volume per 15 µl
2x SYBR Green master mix [a]	7.5 µl
Milli-Q H_2O	4.5 µl
Forward primers	1 µl
Reverse primers	1 µl
Template	1 µl

[a] recipe shown in Table 3.7

3.2.8 Cloning of the full length *LpSH1* gene

3.2.8.1 Primer design for *LpSH1* gene

The coding region of *LpSH1* was identified from the perennial ryegrass transcriptome database. The structure of *LpSH1* gene was determined based on the result of multiple alignments of *LpSH1* and its homologues in other related species. The full length of *LpSH1* is 5-7 kb (Appendix 3.1). The *LpSH1* coding sequence contains an Open Reading Frame (ORF) and two untranslated regions (UTR). With the *LpSH1* cDNA sequence, the intron primers were designed to span the exon boundary. All primer sequences are listed in Appendix 3.1.

3.2.8.2 Amplification for *LpSH1* intron sequences

All primer pairs were tested and selected by using standard PCR with ryegrass genomic DNA as temperate. The genomic DNA was extracted from cv. Glencar following the protocol described in Chapter 2. A total 25 μl of PCR reaction contained 0.2 mM of dNTPs, 1.25 units of TransStart *Taq* DNA Polymerase (TransGen Biotech, China), 1x TransStart *Taq* buffer, 0.2 μM of each primer and 200 ng of genomic DNA. The reaction was incubated at 94°C for 7 min, followed by 35 cycles of 94°C for 30s, 56°C for 30s, and 72°C for 1-2 min, depending on the amplicon length, and finally incubated at 72°C for 10 min. The PCR products were checked on 1% (w/v) agarose gel with TAE buffer at 10 V/cm for 20 min. The selected primer sequences are listed in Table 3.9. The PCR products amplified using the selected primers were separated on 1% (w/v) agarose gel and the target bands were collected for gel purification.

Table 3.9 Selected primer sequences for introns of *LpSH1*

Name	Sequences	Region	Expected size
LpSH1F3	CCACATGCACGCACACTTGCACTAG	5' UTR	200bp
LpSH1R4	GCGAGAATTGTGTTGCAGAAGTTGCAG		
LpSH1F7	CCGGCGGAGCATGTGTGCTAC	Intron 1	2.5kb
LpSH1R6	GTGGTGGTGACTGGATTAAACCTCTCAAG		
LpSH1F9	TGCTGAACATCGTGACCGTTCGTTG	Intron 2	800bp
LpSH1R8	GTCACTTTTGGTTGAGTACATCATTGGCATAC		
LpSH1F12	CCAAATACCGTATGCCAATGATGTACTCAAC	Intron 3	1.5kb
LpSH1R10	GGTTATACGCTGAAGGAACCCGCTG		
LpSH1F14	GCAGCGGGTTCCTTCAGCGTATAAC	Introns 4&5	1.5kb
LpSH1R15	AGGAATTTAATTTTGGCATTATTCAAACCAGTTCCT		
LpSH1I1F1	CGTTGTGTAGAGTGTTGCTATTCTTTATCTGAATC	middle part of intron 1	500bp
LpSH1I1R1	TGTGATCTTCGATGTATAGAGAGAGAGATCGAG		

3.2.8.3 TA-Cloning

The purified PCR products were ligated to the pMD18-T Vector (TaKaRa, Japan). Each reaction contained 0.1 μl of pMD18-T Vector, 2.5 μl of Solution I (with the vector kit, TaKaRa, Japan), and 2.4 μl of each purification product. The reaction was mixed gently and incubated at 16°C overnight.

The following day, the ligation products were transformed into competent cells prepared from *E. coli* DH5α. 2.5 μl of the ligation reaction was added into 100 μl of competent cells (made by Yantai lab staff). The tubes were gently mixed by flicking the bottom of the tube a few times and the solution then incubated on ice for at least 30 min. The tubes were put in a 42°C water bath for exactly 90s, and then incubated on ice for 3 min. About 0.8 ml of LB medium without antibiotic was added into the tube and the solution was incubated with shaking at 160-180 rpm at 37°C for 1 h. The tubes were spun at 13000*g* for 1 min. The supernatant was removed until about 200 μl remained. The remaining supernatant was then used to resuspend the cell pellet gently and this was spread on LB media plates with ampicillin (100 μg/ml). After the plates dried, they were moved into a 37°C incubator overnight.

3.2.8.4 Plasmid DNA isolation

Plasmid DNA was isolated using the alkaline-method (Yantai lab protocol), based on Birnboim (1988). A single isolated colony picked from an LB plate was inoculated into a test tube containing 3-5 ml LB containing ampicillin. The test tubes were incubated overnight at

37°C with shaking at 200 rpm. The bacteria pellet was collected by spinning at 13000*g* for 1 min at room temperature. The bacterial pellet was suspended completely by vortexing in a 160 µl of Solution I (pH 8.0) [50 mM glucose, 10 mM EDTA, 25 mM Tris]. 320 µl of Solution II [400 mM sodium hydroxide and 2% sodium dodecyl sulfate (SDS)], was added and mixed by inverting gently. Then 240 µl of Solution III (pH4.8) [5M potassium acetate], was added and the contents mixed by inverting the tube gently. Approximately 700 µl chloroform was added to the tubes and mixed vigorously. The tubes were centrifuged for 15 min at 13000*g*. The supernatant then transferred to a new 1.7 ml tube without disturbing the white pellet. An equal volume of pre-chilled isopropanol was added into the supernatant and the contents were mixed well by inverting. The mixture was centrifuged at 13000*g* for 15 min at 4°C, and the supernatant discarded. The DNA pellet was washed twice with 70% ethanol, and the pellet left to air dry at room temperature. The pellet was resuspended in 50 µl of sterile H_2O and stored at -20°C.

3.2.8.5 Sequencing and assembling

All isolated plasmids containing amplicons of *LpSH1* were sent to Sangon Biotech company (China) for sequencing. The primers for *LpSH1* introns were designed by using the sequences of the flanking exon regions. After sequencing, the exon region sequences were used for splicing, as the flanks of amplicons overlap each exon. The full length *LpSH1* gene was assembled by hand using the exon regions from the transcriptome database, aligning these with the flanks of the amplicons (i.e. the exon regions), and including the sequences of the introns.

3.3 Results

3.3.1 RNA extraction and cDNA synthesis

Total RNA was extracted from tissue enriched for abscission layers collected at different developmental stages. RNA quality is one of the most important determinants of the subsequent RT-qPCR results. The total RNA was checked on 1% (w/v) agarose gels with TAE buffer at 10 V/cm for RNA integrity and quality (Figure 3.3) and assessed by a NanoDrop spectrophotometer for purity and quantity (Appendix 3.2). In Figure 3.3, it is clear that there are two sharp bands of the 28S and 18S RNA, and the ratio of 28S RNA to 18S RNA is about 2 to1, which means that the total RNA is of good quality. A A260/280 ratio of

over 1.8 is generally indicative of pure RNA. In this study, the A260/280 ratio of RNA products was in the range of 1.9-2.2, suggesting that the RNA products were of good purity. The total RNA was directly used for cDNA synthesis and the diluted cDNA products were then checked and normalised by reference genes using RT-qPCR.

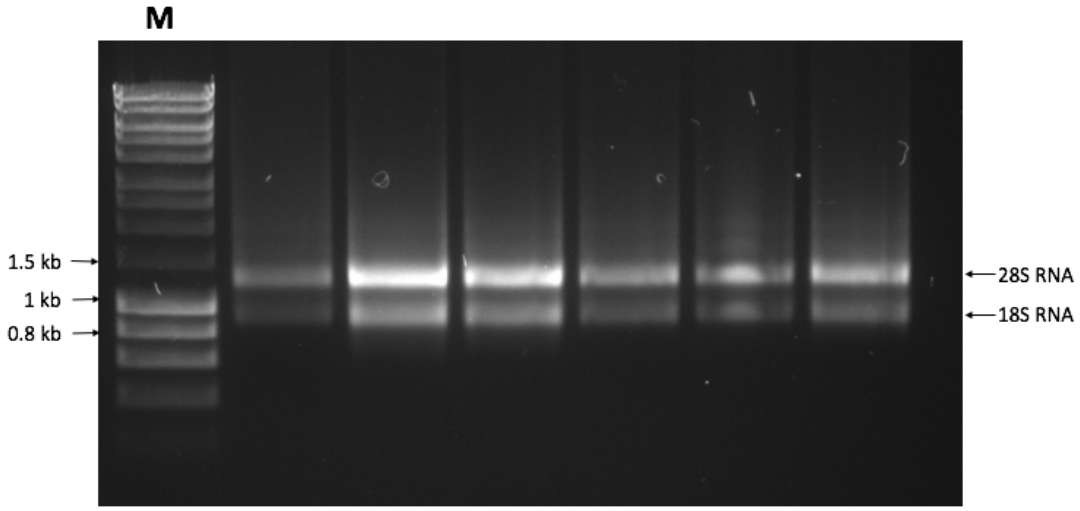


Figure 3.3 RNA integrity check. Total RNA isolated from ryegrass spikelets using the Qiagen RNeasy Plant Mini kit was checked on a 1% (w/v) agarose gel with TAE buffer at 10 V/cm. The bands for 28S and 18S RNA are indicated. M: HyperLadder I.

3.3.2 Primer pair selection and PCR optimisation for gene expression study

The PCR efficiency of different primer pairs for the same gene could be different, as the PCR products may have different amplicon lengths, sequences and annealing temperature. These facts affect the C_t value and product specificity during qPCR. Figure 3.4 shows the RT-qPCR performance of the reference genes with different primer pairs. From the melting curve analysis (Figure 3.4 A), and the products amplified by primers, LpEF-F1R, LpEF-F3R and LpGap-FR1, show sharp, single peaks, respectively, suggesting that these PCR products are specific. The quantitation analysis (Figure 3.4 B) shows a similar C_t value of LpEF-F1R and LpEF-F3R, so both pairs could be used for the gene expression study. Compared with these three pairs, other primer pairs generated wider peaks, higher C_t numbers and primer dimers, suggesting that these primer pairs have low amplification efficiency (Figure 3.4). The PCR performance could be significantly changed if even only one of the primers was replaced. Therefore, the quantitation analyses and melting curve analyses were compared for different primer combinations for each target gene. The primer pairs with the highest amplification efficiency were selected for gene expression study and are listed in Table 3.10.

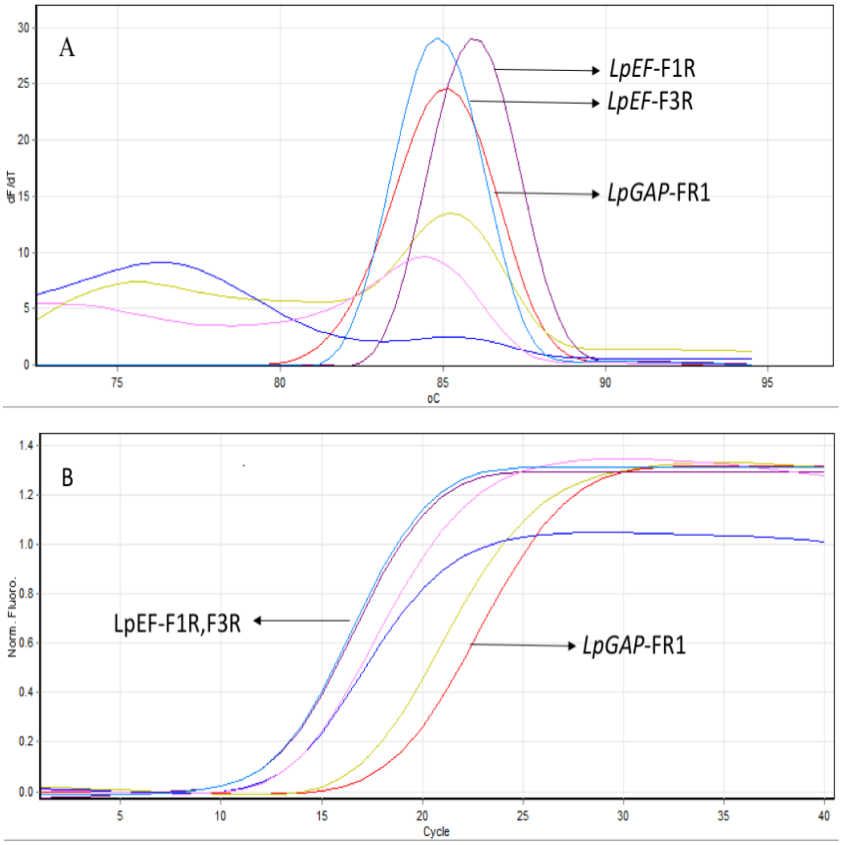


Figure 3.4 The RT-qPCR performance with different ryegrass reference gene primer pairs. A: Melting Curve Analysis; B: Quantitation Analysis. Primer pairs LpEF-F1R, LpEF-F3R and LpGap-FR1 are indicated. Yellow: LpGAP-FR2; Pink: LpEFF2R; Dark blue: LpGAP-FR3 (refer to Table 3.5 for sequences).

Table 3.10 Reference and target gene primer sequences selected for the gene expression study

Name		Sequences	Amplicon size (bp)
LpSH1	F2	AAGAGGCAGCGGGTTCCTTCAG	189
	R1	GATGGCCTCGTCCAGCTTCTT	
LpLG1	F1	TGGTGAAGGAAATGCAGGTGGATC	210
	R1	CAGGCTGTTCTTGCTGCACTTG	
LpSH4	F1	ACTACCGCAAGGGCAACTGGAC	190-220
	R1	GTCGTTGCACTGGTTCTGGCTG	
LpqSH1b	F2	GCTACAAGCAATACTACCAGCAGCT	221
	R1	CCCGAAGTTCGCCATCTCTTCCT	
LpQ	F2	AGGAGGAGTTCGTGCACATCCTC	216
	R2	CTTCCCTCCCRTTGAAGCGAATG	
LpqSH1a	F3	GGCAGTACTACCAGCAGGTCCAG	194
	R3	GATGTCTTTGCTCATGCCTTCTTTGAC	
LpSHAT1a	F	CGAGGATTACGAGGACATCAAGCA	210
	R	TCCCATCTGCCGCACTTGTG	
LpGAP	F	AGGAGGTTGCYGTSTTTGGCTG	237
	R1	GTCAGASKTGTAYTCCTTCTCRTTG	
LpEF	F1	CACCCTGGTCAGATCGGCAAC	238
	R	CACCAACAGCAACAGTCTGCCT	

3.3.3 Isolation of candidate seed shattering genes from perennial ryegrass

The PCR products for nine candidate seed shattering gene sequences, two for *LpqSH1*, *LpqSH1a* and *LpqSH1b*, respectively, and one each for *LpSH1*, *LpLG1*, *LpSH4*, *LpQ*, *LpSHAT1a, LpSHAT1b* and *LpWRKY*, were amplified using the optimised primer pairs. These candidate gene sequences were validated by sequencing. The sequencing results of all candidate genes were verified with their homologues via BLASTN searching through the database of nucleotide collection and expressed sequences tags on NCBI website. A BLAST algorithm was optimised for more dissimilar sequences. The Neighbour-Joining phylogenetic tree was constructed using all of the newly isolated candidate seed shattering gene sequences and their homologues in related monocot species from the NCBI website. The rooted phylogenetic tree from multiple sequence alignments is presented in Figure 3.5, demonstrating that each candidate gene in ryegrass groups together with its corresponding homologues in other monocots.

The initial phylogenetic tree constructed in early work of this thesis showed that *LpqSH1a* and *LpqSH1b* grouped together with *OsqSH1* in rice. Recently, Yoon et al. (2014) demonstrated that *OsSH5* is highly homologous to *OsqSH1*, and both genes play a role in abscission layer formation and differentiation. Therefore, the realigned phylogenetic tree (Figure 3.5) suggests that *LpqSH1a* is homologous to *OsqSH1* and *LpqSH1b* is homologous to *OsSH5*.

In the phylogenetic tree, *LpSHAT1a* clustered into the group of *SHAT1* with other monocot species, suggesting that *LpSHAT1a* is homologous to *SHAT1*. However, *LpSHAT1b* clustered with two expressed sequences tag from perennial ryegrass which had a close relationship with *LpQ*. Consequently, *LpSHAT1b* was ruled out of the gene list for expression study.

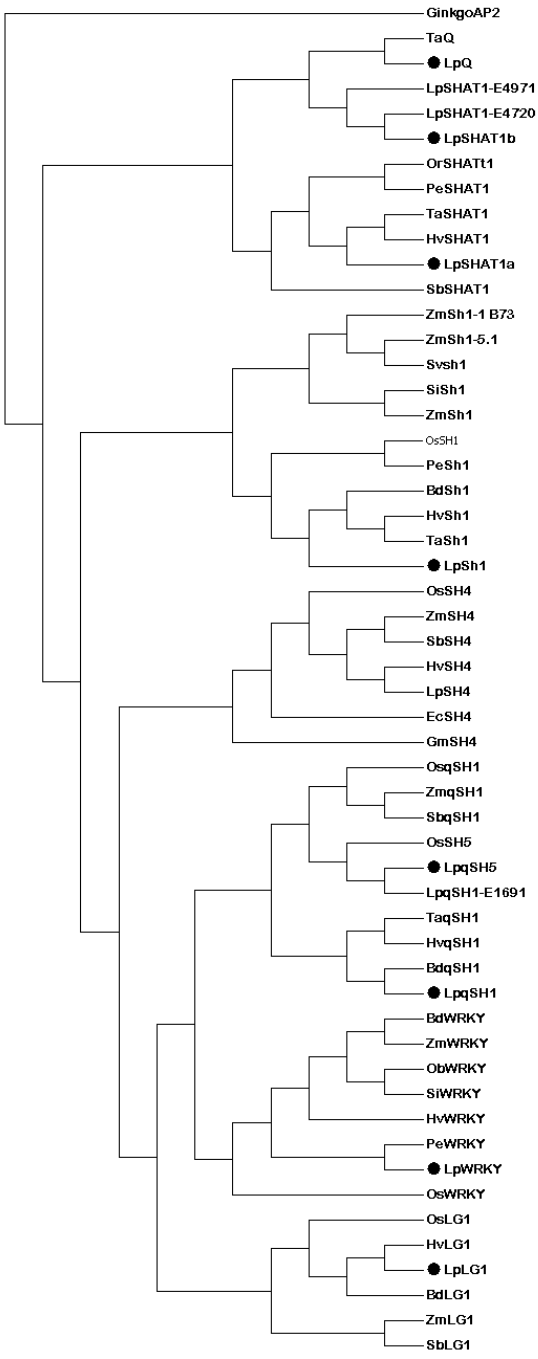


Figure 3.5 Rooted phylogenetic tree of candidate target seed shattering homologues in monocots. All homologues from perennial ryegrass are marked [·]. The phylogenetic tree was rooted using Ginkgo *AP2*. Each homologue was renamed for easy reference in this thesis. Refer to Table 3.4 for accession numbers. *Ta, Triticum aestivum; Os, Oryza sativa; Lp, Lolium perenne; Or, Oryza rufipogon; Pe, Phyllostachys edulis; Hv, Hordeum vulgare; Sb, Sorghum bicolor; Zm, Zea mays; Sv, Sorghum virgatum; Si, Setaria italic; Bd, Brachypodium distachyon; Ec, Echinochloa crus-pavonis.*

3.3.4 Relative expression of candidate seed shattering genes in perennial ryegrass

It is possible that each of the candidate genes plays a role in the seed shattering process in perennial ryegrass. To determine this, quantitative expression analysis of the candidate genes in the abscission layer at different spikelet/seed developmental stages was carried out using RT-qPCR. Using *GAPDH* (*GAP*) and *Elongation Factor* (*EF*) as reference genes, the relative expression levels of candidate shattering genes in three biological replicates are presented in Figures 3.6 to 3.8. From the relative expression study, six genes, *LpqSH1*, *LpSH5*, *LpLG1*, *LpSH4*, *LpSH1*, and *LpSHAT1* showed changes in expression level at different developmental stages of the abscission layer and vegetative tissues (Figures 3.6-3.8). Additionally, the overall patterns of relative expression level in the three biological replicates are presented in Figures 3.9 to 3.11.

At the early spike/spikelet developmental stages, *LpLG1* had an extremely high expression level compared with during seed development in the first and third biological replicates. Its expression level was also high in seedling leaves (Figures 3.6 and 3.8). Early spike development and seedling leaves were not included in the second biological replicate (Figure 3.7)

The *LpqSH1* (originally *LpqSHa*) and *LpSH5* (originally *LpqSHb*) genes expressed at a high level in spikes of 1-2 cm long, then their expression level declined till spikes reached 10 cm in length. High expression levels were also detected in late seed developmental stages and node in the first and the third biological replicates (Figures 3.6 and 3.8). However, in replicate 2, the expression of both genes was low at 18 and 21 daa (Figure 3.7). For *LpSH4* and *LpSHAT1*, there expression level markedly increased when spikes of 1-2 cm to 4-8 cm long in the first biological replicate (Figure 3.6). *LpSH4* had an increased expression in the late seed developmental stages in the three biological replicates (Figures 3.6-3.8) *LpWRKY* was highly expressed in flag leaves and mid-seed developmental stages (Figures 3.6-3.8).

LpSH1 was highly expressed post anthesis, in the first biological replicate, from early stages of seed development to mature seed (Figure 3.6). During spike development, the relative expression level of *LpSH1* was low, but increased markedly from 0 to 5 daa, being around 1400-fold higher than the baseline expression. Another increase in expression occurred from 10 to 15 daa, being about 3-fold higher than that at 5 daa (Figure 3.6). A final increase in expression occurred at 20 daa, being 1.5-fold higher than 15 daa, which was 6000-fold higher

than the baseline expression (Figure 3.6). The second and third biological replicates shared a similar expression pattern: two peaks were detected after anthesis, at 4 and 10 daa (Figure 3.7) and at 6 and 14 daa (Figure 3.8). After 14 daa in both biological replicates, the expression sharply dropped.

From the overall relative expression patterns across the three biological replicates, the expression pattern of *LpSH1* presented the most noticeable changes (Figures 3.9-3.11). The transcript level of *LpSH1* increased during spike/seed development. At the early stage of AL development, *LpSH1* expression was low compared to some of the other candidate genes. However, at later stages, *LpSH1* expression was much greater than that of the other candidate genes (Figure 3.9-3.11).

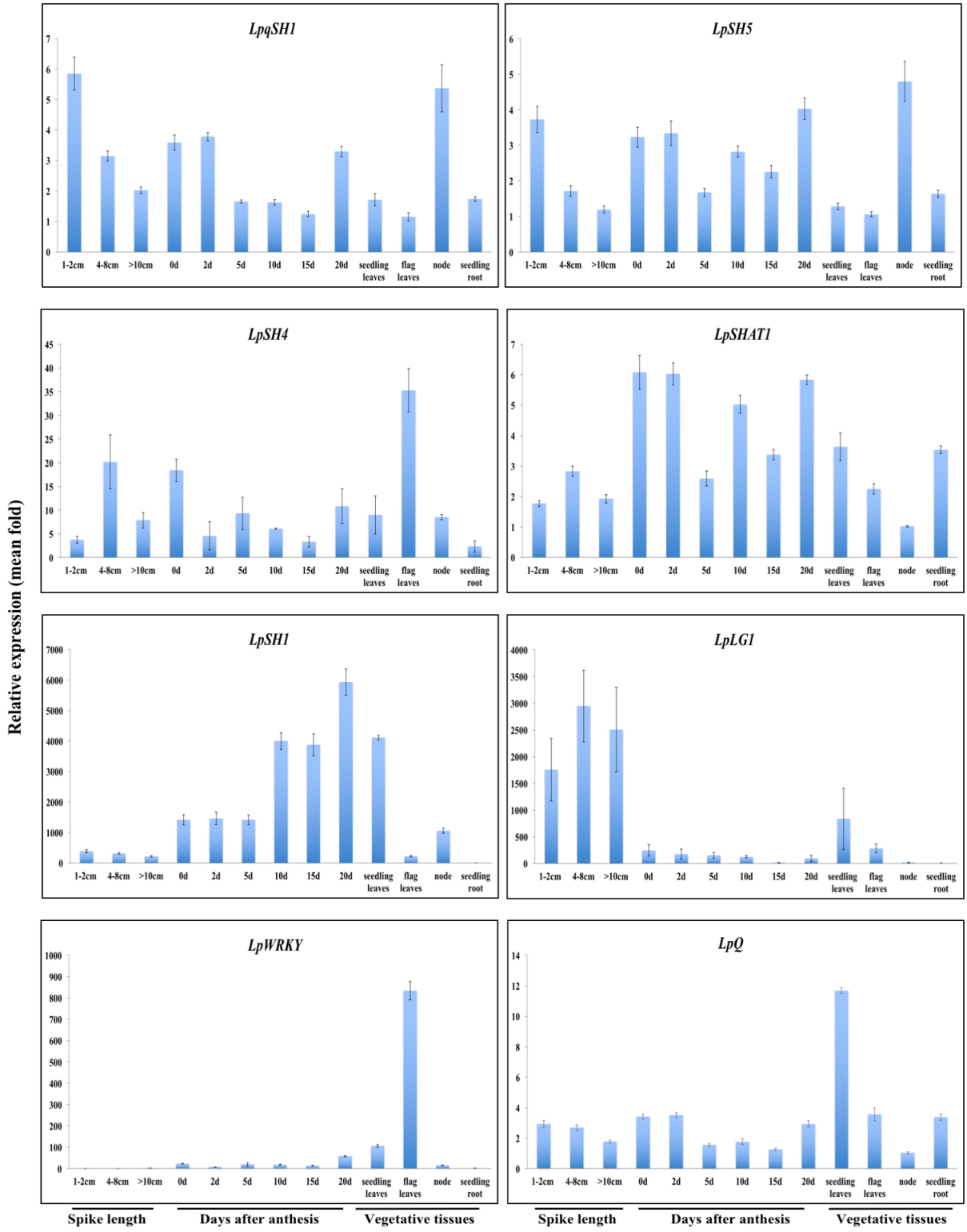


Figure 3.6 Relative expression of putative *LpqSH1*, *LpSH5*, *LpSH4*, *LpSHAT1*, *LpSH1* *LpLG1*, *LpWRKY*, and *LpQ* in perennial ryegrass cv. RI009 (first biological replicate). The plant materials were collected the glasshouse at the University of Canterbury, New Zealand, from October to December in 2014.

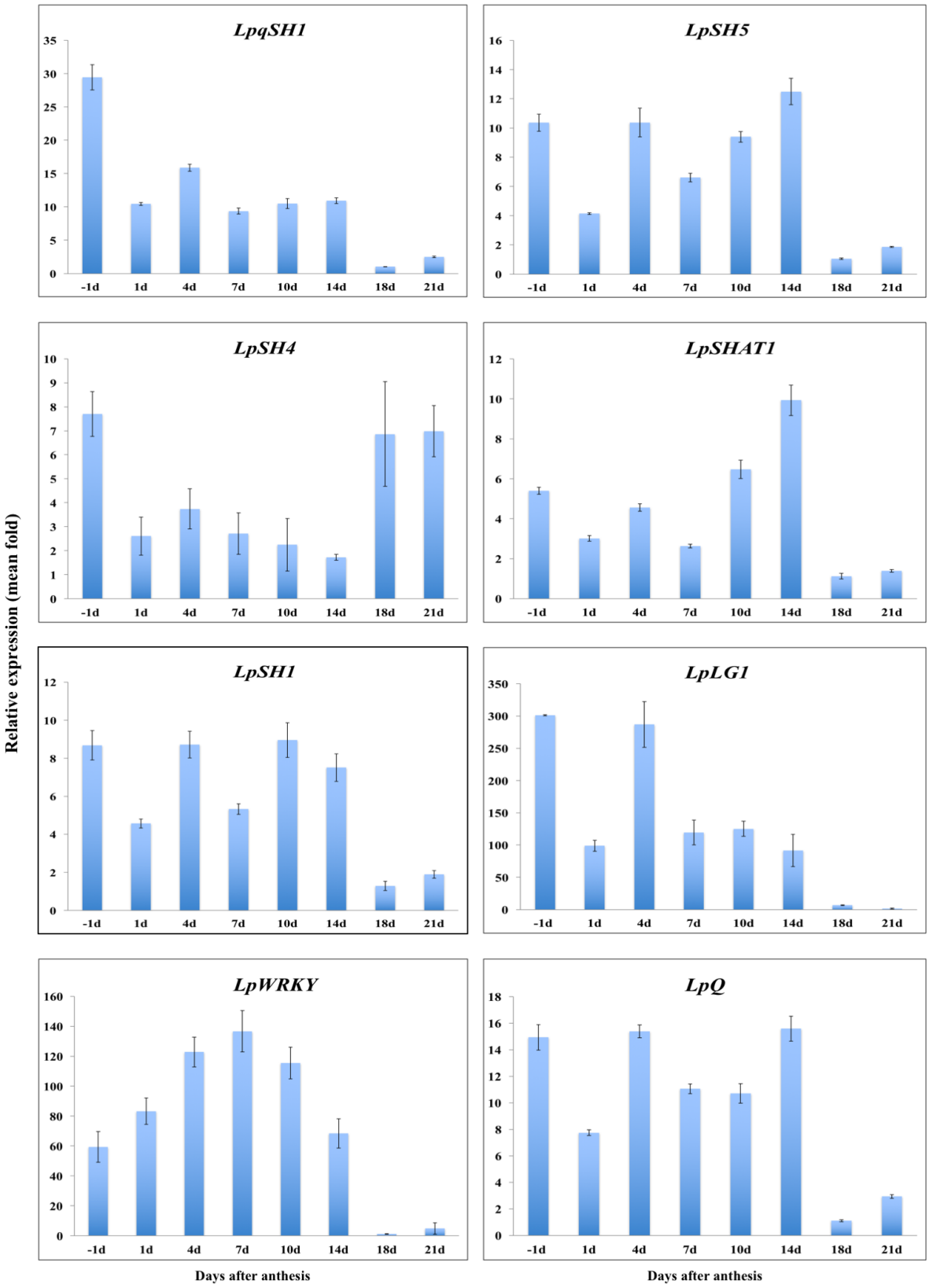


Figure 3.7 Relative expression of putative *LpqSH1*, *LpSH5*, *LpSH4*, *LpSHAT1*, *LpSH1* *LpLG1*, *LpWRKY*, and *LpQ* in perennial ryegrass cv. Nui (second biological replicate). The plant materials were collected from a field located outside of Christchurch, New Zealand, from October to December in 2014.

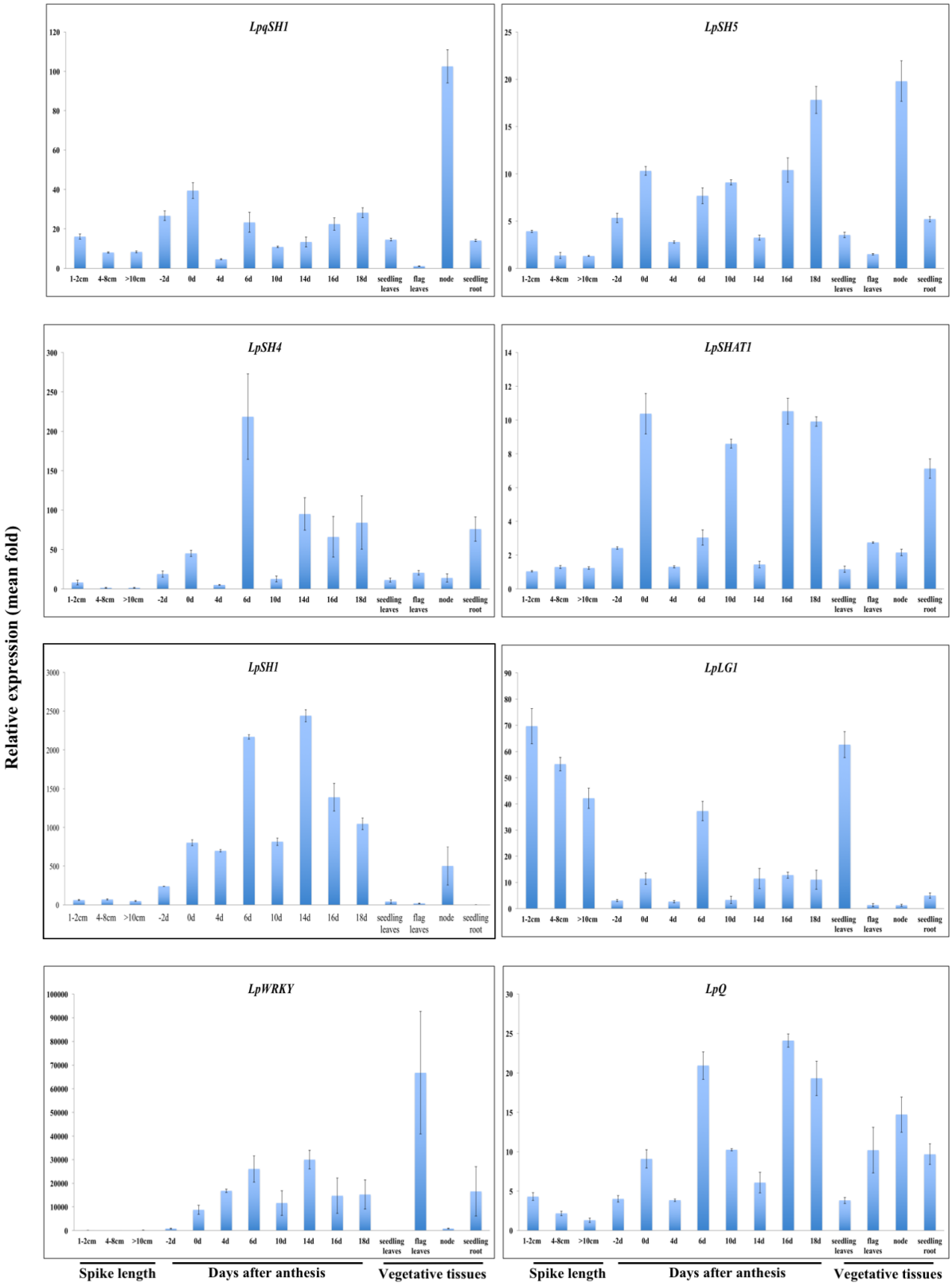


Figure 3.8 Relative expression of putative *LpqSH1*, *LpSH5*, *LpSH4*, *LpSHAT1*, *LpSH1* *LpLG1*, *LpWRKY*, and *LpQ* in perennial ryegrass cv. Nui (third biological replicate). The plant materials were collected from the glasshouse in Yantai University, China, from May to June in 2015.

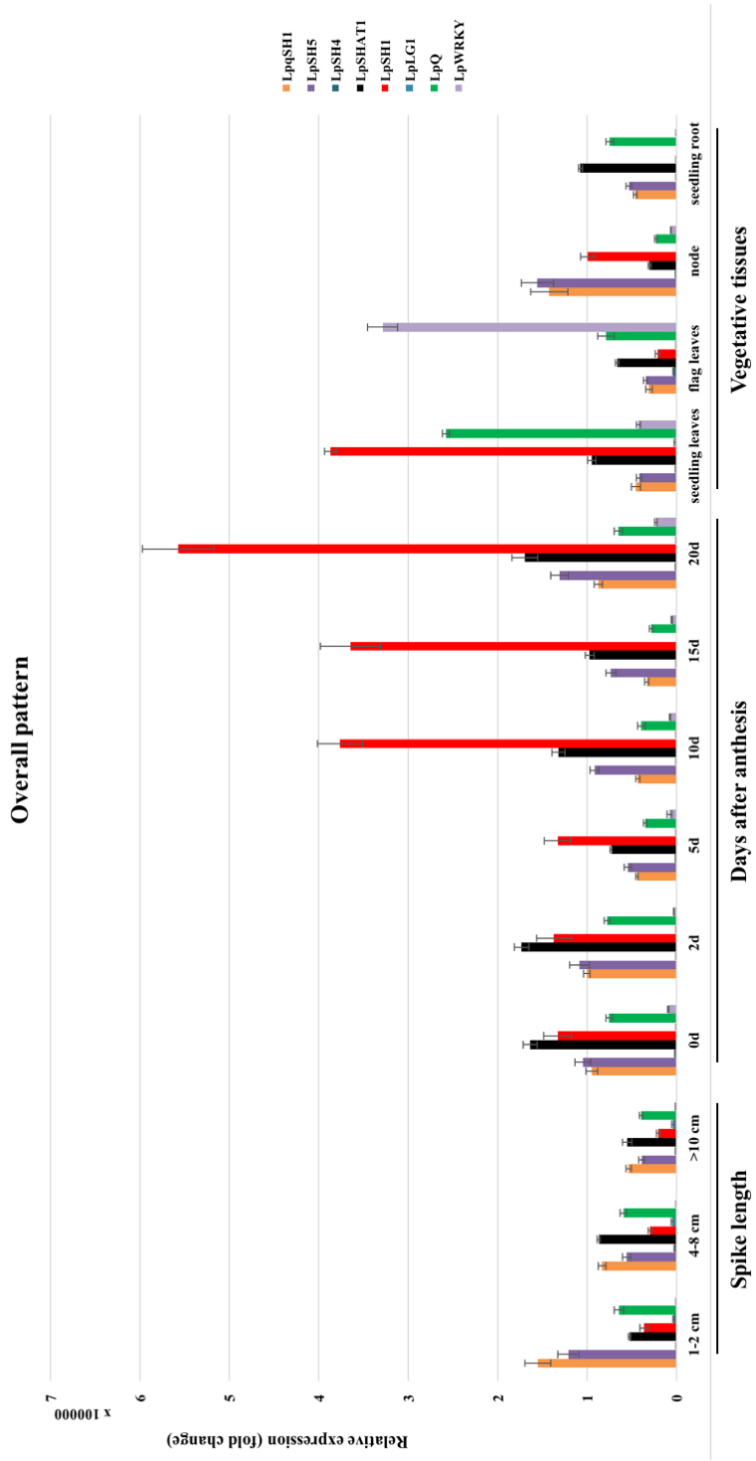


Figure 3.9 The overall pattern of relative expression level of candidate seed shattering genes in first biological replicate. The details are described in Figure 3.6.

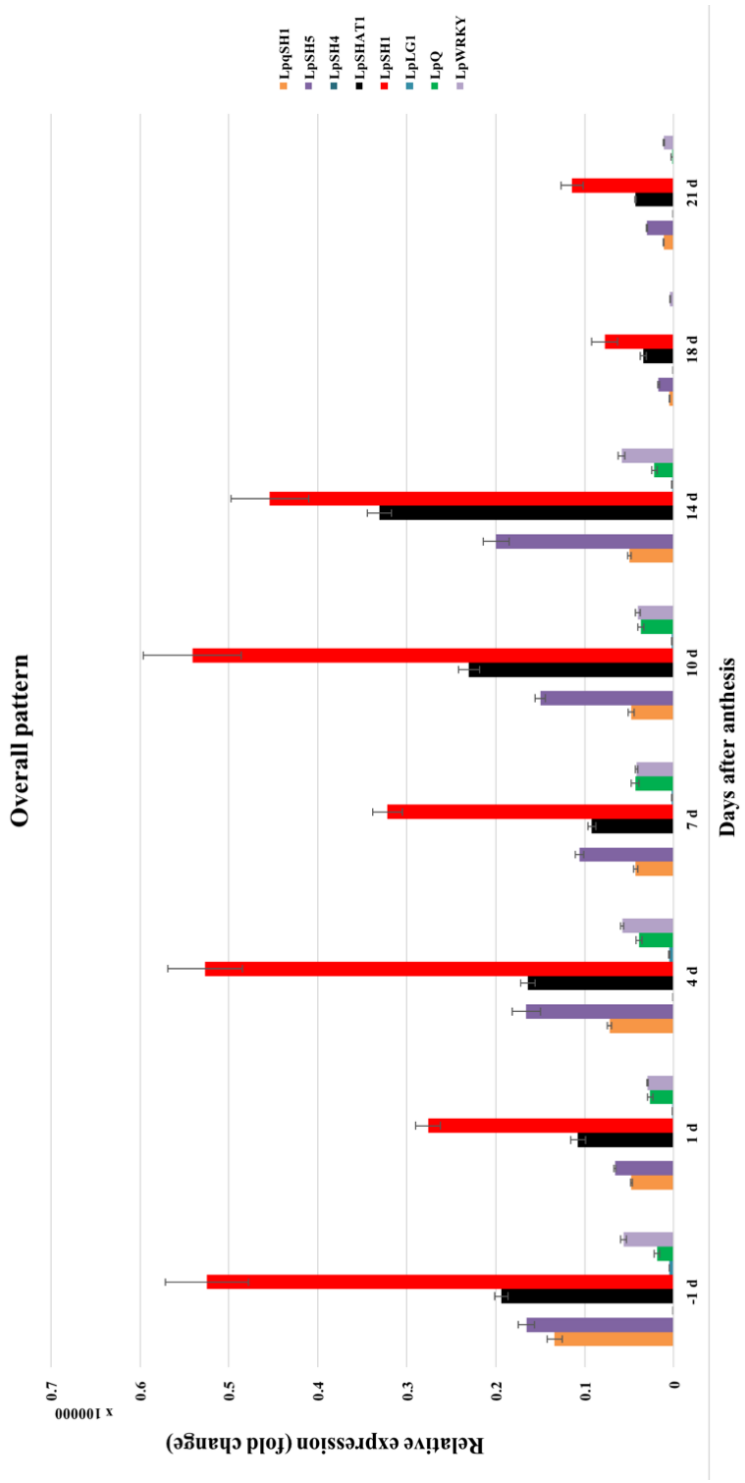


Figure 3.10 The overall pattern of relative expression level of candidate seed shattering genes in second biological replicate. The details are described in Figure 3.7.

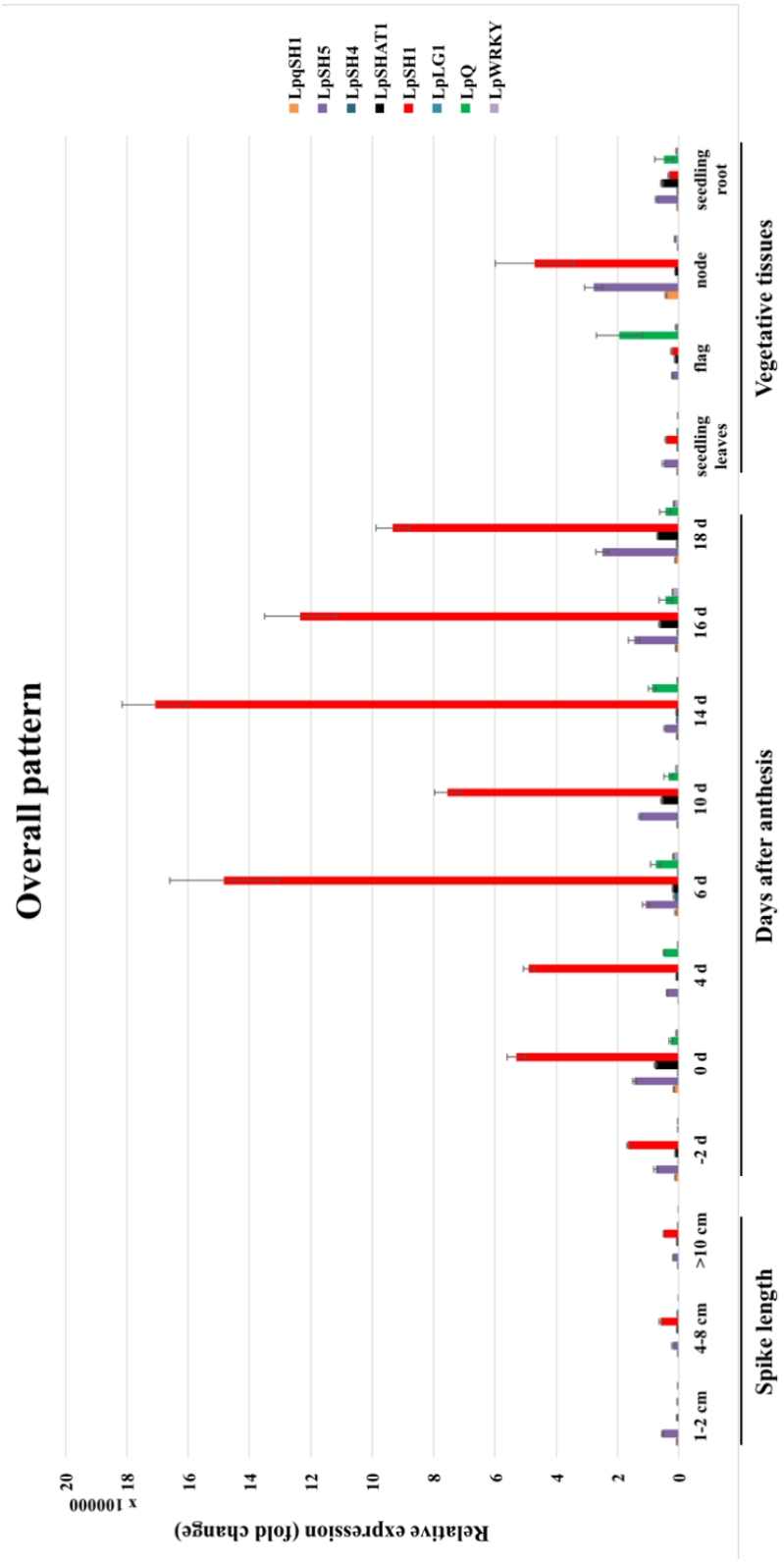


Figure 3.11 The overall pattern of relative expression level of candidate seed shattering genes in third biological replicate. The details are described in Figure 3.8.

3.3.5 Determination of the full length *LpSH1* gene

The ORF of *LpSH1* was identified from the ryegrass transcriptome database, and its intron sequences were amplified from genomic DNA and sequenced using TA-cloning. The full length of the *LpSH1* sequence is attached in Appendix 3.3. The *LpSH1* ORF contains five introns and six exons, and the lengths from the first to fifth introns are about 2427, 722, 1792, 98, and 436 bp, respectively. The structure of the full length *LpSH1* gene is displayed in Figure 3.12.

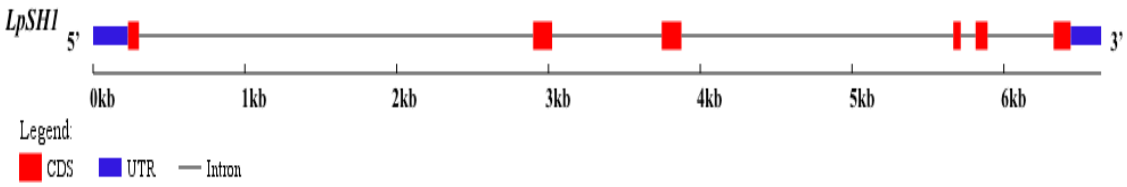


Figure 3.12 Gene structure of *LpSH1*

The length of the *LpSH1* coding region is 564 bp, encoding a polypeptide of 187 amino acids. The protein sequence contains two conserved domains, including a C2C2 zinc finger domain from amino acid positions 11 to 47 and a helix-loop-helix YABBY domain from amino acid positions 108 to 156 (Figure 3.13). The phylogenetic tree was constructed using the result of multiple alignments of the amino acid sequence of LpSH1 and its homologues from related monocot species, showing LpSH1 is closely related to other YAB2 subfamily members. This suggests that LpSH1 belongs to the YAB2 subfamily (Figure 3.14). The amino acid sequence of LpSH1 has 93% identity with the predicted protein HvYABBY (NCBI accession No. BAJ89435.1) and 95% identity with TaYABBY2 (NCBI accession No. ABW80974). From multiple alignments of amino acid sequences from the YAB2 subfamily, the two conserved domains have high similarity (Figure 3.13).

```
LpSH1       ..........MSAQIAP.AEHVCYVHCNFCNTILAVSVPSNSMLNIVTVR     39
HvYABBY2    ..........MSAQIAP.pEHVCYVHCNFCNTILAVSVPSNSMLNmVTVR     39
ZmYABBY2    myhtqgegmssaAQIAPvpEHVCYVHCNFCNTILAVSVPghSMLNmVTVR     50
SiYABBY2    .........msapQIvPalEHVCYVHCNFCNTILAVSVPSNSMLNIVTVR     41
ObYABBY2    ..........MSAQIvPapdHVCYVHCNFCNTILAVSVPSNSMLNIVTVR     40
TaYABBY2    ..........MSAQIAP.pEHVCYVrCNFCNTILAVSVPSNSMLNIVTVR     39
OsYABBY2    ..........MSAQIvPapEHVCYVHCNFCNTIfAVSVPSNSMLNIVTVR     40

LpSH1       CGHCTSLLSVNLRGLIQSPPPVQDH....SQENLKAHNISFR.GNYPDYS     84
HvYABBY2    CGHCTSLLSVNLRGLIQSPlPVQDH....SQENfKAqNISFh.GNYPDYg     84
ZmYABBY2    CGHCTSLLSVNLRGLIQSl.PVvqn..hySQEhfKvqNfSFt.eNYPeYa     96
SiYABBY2    CGHCTSLLSVNLRGLIQSl.PVQnHlcllvQENLKvHNfSFsteNYceYa     90
ObYABBY2    CGHCTSLLSVNLRGLvQSl.PaeDH....lQENLKmHNISFR.eNYseYg     84
TaYABBY2    CGHCTSLLSVNLRGLIQSP.PVQDH....SQENfKAHNISFR.GNYPDYg     83
OsYABBY2    CGHCTSLLSVNLRGLvQal.PaeDH....lQdNLKmHNmSFR.eNYseYg     84

LpSH1       ...SSSKY.RMPMMYSTKSDPEHMLHMRPATEKRQRVPSAYNRFIKEEIR    130
HvYABBY2    ...tSSKY.RMPMMfSTKSDqEHMLHMRPApEKRQRVPSAYNRFIKEEIR    130
ZmYABBY2    appSSSrY.RMPtMlSaKgDldHMLHvRa.pEKRQRVPSAYNRFIKEEIR    144
SiYABBY2    ..pSSSKY.RMPtMfSTKgDqdHMLHvRa.pEKRQRVPSAYNRFIKEEIR    136
ObYABBY2    ...SSSrYgRvPMMfS.KnDtEHMLHvRP.pEKRQRVPSAYNRFIKEEIR    129
TaYABBY2    ...tSSKY.RMPMMfSTKSDqEHMLHMRP.pEKRQRVPSAYNRFIKEEIR    128
OsYABBY2    ...SSSrYgRvPMMfS.KnDtEHMLHvRP.pEKRQRVPSAYNRFIKEEIR    129

LpSH1       RIKTNNPDISHREAFSTAAKN...WAHFPNIHFGLGSNESSKKLDEAIAA    177
HvYABBY2    RIKaNNPDISHREAFSTAAKN...WAHFPNIHFGLGSNESSKKLDEtIAt    177
ZmYABBY2    RIKaNNPDISHREAFSTAAKNkymWAHFPNIHFGLGpyESSnKLDEtIgA    194
SiYABBY2    RIKasNPaISHREAFSTAAKN...WAHFPNIHFGLGplESSKKLDEAIgA    183
ObYABBY2    RIKaNNPDISHREAFSTAAKN...WAHFPNIHFGLGphESSKKLDEAIAA    176
TaYABBY2    RIKTNNPDISHREAFSTAAKN...WAHFPNIHFGLGSNESSKKLDEAIAA    175
OsYABBY2    RIKaNNPDISHREAFSTAAKN...WAHFPNIHFGLGShESSKKLDEAIgA    176

LpSH1       P.IPQKVQGLY                                           187
HvYABBY2    P.IPQKVQGLY                                           187
ZmYABBY2    tghPrKiQdpY                                           205
SiYABBY2    aghPhKVQdLY                                           194
ObYABBY2    P.gPQKVQrLY                                           186
TaYABBY2    P.IPQKVQGLY                                           185
OsYABBY2    P.sPQKVQrLY                                           186
```

Figure 3.13 Amino acid sequence alignment of LpSH1 and homologous gene proteins. A zinc-finger domain is indicated with the shaded box. The YABBY domains are underlined and in bold in LpSH1. The protein accessions are as follow, HvYABBY2 (*Hordeum vulgare*, BAJ89435.1); SiYABBY2 (*Setaria italica* XP_004982272.1); ObYABBY2 (*Oryza brachyantha* XP_006650352.1); TaYABBY2 (*Triticum aestivum* ABW80974.1); OsYABBY2 (*Oryza sativa* XP_015628574.1); ZmYABBY2 (*Zea mays* XP_008666788.2).

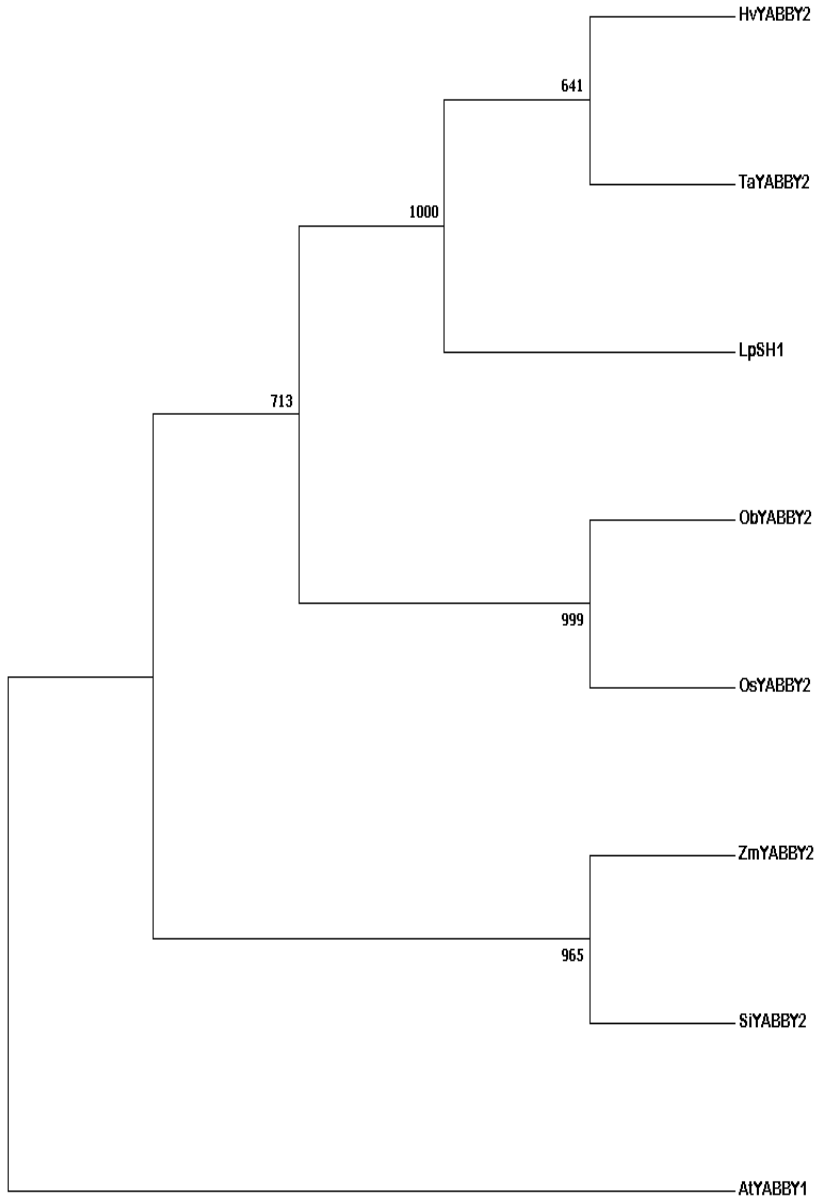


Figure 3.14 Rooted phylogenetic tree of the YABBY2 sub-family in monocots. The tree is rooted by a family member in arabidopsis. The protein accessions are as follows: HvYABBY2 (*Hordeum vulgare*, BAJ89435.1); SiYABBY2 (*Setaria italica* XP_004982272.1); ObYABBY2 (*Oryza brachyantha* XP_006650352.1); TaYABBY2 (*Triticum aestivum* ABW80974.1); OsYABBY2 (*Oryza sativa* XP_015628574.1); ZmYABBY2 (*Zea mays* XP_008666788.2); AtYABBY1 (*Arabidopsis thaliana* AAD33715.1). Node values are the number of times the split was identified in 1000 bootstrap replicates.

3.4 Discussion

To investigate genes involved in the seed shattering trait in perennial ryegrass, a comparative genomics approach was undertaken. The eight seed shattering candidate genes were selected based on the published studies in cereal crops. Using these genes as query sequences, seed shattering homologous genes were identified from a perennial ryegrass transcriptome database recently created in this laboratory. The result of phylogenetic analysis (Figure 3.5) showed that all candidate shattering genes detected from the transcriptome database were grouped with their corresponding homologues. Comparative phylogenetic analyses have been reported as a way to detect homologous and their functions in target species, in particular for some species with little genomic information, as the case in perennial ryegrass (Jones et al., 2002; Morrell et al., 2012; Song et al., 2012). The perennial ryegrass transcriptome used in this work is generated from a pool of combined RNA samples extracted from multiple developmental stages of leaves, flower spikes and seeds (Roche et al., 2016; Guo et al., 2017). In this work, all of the candidate seed shattering homologues were successfully identified from this transcriptome database, suggesting that perennial ryegrass homologues of these candidate genes potentially share similar functions to those in rice and wheat in terms of controlling seed shattering.

From the results of gene expression study, *LpLG1* has a high transcript level at early stages before anthesis and then expressed at low level after anthesis. This coincides with the function of *OsLG1* in rice, which plays a role in panicle architecture (Ishii et al., 2013). The putative *LpLG1* also expressed during panicle development in perennial ryegrass, suggesting that *LpLG1* probably has a similar role to that in rice. From another aspect, the fact that *LpLG1* can be detected at the early spike developmental stages also supports the study of Elgersma et al. (1988) that the abscission layer has differentiated and formed before heading. This suggests that AL formation in ryegrass could be at similar developmental stages with that in rice. However, no study showing the precise developmental stages of abscission layer formation in perennial ryegrass has been conducted to date. Therefore, to detect the expression level of these candidate seed shattering genes in abscission layer, plant material almost spanning the whole developmental stages, from young spikes to mature seed, was used in this work.

Using QTL analysis, *qSH1* was determined to an play important role in seed shattering trait in rice (Konishi et al., 2006). According to the DNA sequence of *OsqSH1*, two DNA sequences were identified through BLAST searching the perennial ryegrass transcriptome database. Both sequences were highly homologous to *OsqSH1* in rice, and were named *LpqSHa* and *LpqSHb*.More resently, Yoon et al. (2014) identified another shattering gene in rice, *OsSH5*, which is highly homologous to *OsqSH1*. Both genes play a role in the seed shattering process in rice by promoting abscission layer development and preventing lignin biosynthesis. From the phylogenetic analysis, *LpqSHa* and *LpqSHb* are homologous to *OsqSH1* and *OsSH5*, respectively, and were renamed *LpqSH1* and *LpSH5*. This suggests that BLAST searching through transcriptome database to identify homologous genes is an ideal method to facilitate the transfer of genetic information from well-studied cereal species to perennial ryegrass.

Based on previous studies in monocot crop species, *SHAT1*, *SH4*, *qSH1*, and *SH5* are considered to take part in abscission layer formation and differentiation in rice (Konishi et al., 2006; Li et al., 2006; Zhou et al., 2012; Yoon et al., 2014). Zhou et al. (2012) constructed a genetic model of abscission layer development in rice, in which *SH4* plays a role upstream of *SHAT1* and *qSH1*, and then *qSH1* (feedback) regulates *SHAT1* and *SH4* at sp8l to control abscission layer development (Figure 1.3). The result of expression study in this work showed that when the length of spike was between 1-2 cm, *LpqSH1* has a higher transcript level than other candidate genes, and *LpSH4* and *LpSHAT1* expression increase from spikes length of 1-2 cm to 4-8 cm (Figures 3.6 and 3.8). This indicated that *LpqSH1* probably plays a role at upstream of *LpSH4* and *LpSHAT1* in perennial ryegrass as well. Yoon et al. (2017) reported that the KNOX protein, OSH15, interacts with qSH1 and SH5 to induce seed shattering in rice. The authors also generated *OSH15* RNAi transgenic plant, indicating that the expression level of qSH1 and SH5 were not changed by OSH15, but the expression of *SH4* was reduced significantly in the RNAi plants. This result suggested that *qSH1* and *SH5* play a role at upstream of *SH4* and then promotes abscission layer development. Our results are consistent with the genetic model or regulatory network specifying abscission layer development drawn by Yoon et al. (2017) and Zhou et al. (2012). In addition, according to Zhang et al. (2013), *TaqSH1* could play a role in grain/seed abscission process in both monocots and eudicots. Taken together, *LpqSH1* probably have potential application in the genetic mechanism of seed shattering in perennial ryegrass.

From the expression pattern of *LpSH4*, the relative expression level increased at late seed development stages (Figure 3.7), suggesting that this gene have a role in activation of abscission process in ryegrass. Coincidently, study of Lin et al. (2007) indicated that *SHA1*, allelic to *SH4*, takes part in the degradation of the cell wall in the abscission layer.

In addition to the genes discussed above, *LpSH1* shows interesting changes in expression level. The transcript level of *LpSH1* in the abscission layer increased with spike/seed development. Especially at about two weeks after flowering, the *LpSH1* gene expressed significantly greater than at any other stage, and the expression level then slightly reduced (Figure 3.6). This may indicate that *LpSH1* has a function during seed developmental stage, and could play a role in abscission layer separation. However, Lin et al. (2012) suggested that *SH1* in sorghum takes part in abscission layer formation, and some non-shattering variants with repressed *SH1* gene expression do not develop an abscission layer, based on microscopic examination during flowering. This evidence indicates that *SH1* in sorghum probably expresses early in development to promote differentiation and formation of the abscission layer. In contrast, our results showed that *LpSH1* expressed strongly during seed developmental stage. These two contrasting observations suggested that *SH1* probably not only plays a role in abscission layer formation, but also continuously maintains abscission activation in perennial ryegrass.

Lin et al. (2012) demonstrated that *SH1* homologues in other monocot species including rice, sorghum and maize, play a role in the grain abscission process. *SH1* is a YABBY-like gene, belonging to the YAB2 subfamily (Lin et al., 2012). This is a coincidence with a phylogenetic analysis of amino acid sequences of LpSH1 and other homologues, indicating YABBY-like gene. The YABBY gene family consists of five clades, including CRABS CLAW (CRC), INNER NO OUTER (INO), FILAMENTOUS FLOWER (FIL), YABBY2 (YAB2), and YABBY3 (YAB3). This gene family regulates complex patterning and growth decisions relating to lateral organ development (Bowman, 2000). *OsSh1* in rice, *Sh1* in sorghum, and *ZmSh1-5.1* and *ZmSh1-1* in maize, all of these homologues could play a role in the formation of AL (Lin et al., 2012), rather than having a major role in polarity regulation in lateral organs (Toriba et al., 2007). Lin et al. (2012) indicated that the *SH1* homologues identified in different cereals have a parallel selection in domestication history for seed shattering trait. Except *SH1* gene, other YABBY genes such as *fasciated* (*FAS*) in tomato (Cong et al., 2008), *DROOPING LEAF* (*DL*) in rice (Ohmori et al., 2008), and *OsYAB1* in

rice (Jang et al., 2004), have been illustrated to play a crucial role in unique functions. Combining previous studies and the gene expression study, it is possible that *LpSH1* plays a critical role in the seed shattering process in ryegrass. Consequently, *LpSH1* was selected as the target gene for the mutation screening experiments described in Chapter 5.

In addition, a transgenic study is being conducted in this work to determine the function of *LpSH1*. Three expression vectors were constructed, including two for over-expression of *LpSH1* and one using RNAi. As the experiment is still on-going, a brief description of experiment of the process is provided in Appendix 3.4.

4 Chapter 4: Optimisation of method for mutation screening

4.1 Introduction

High resolution melting (HRM) analysis is a relatively new technique based on PCR, but was first described in 2003 (Gundry et al., 2003; Wittwer et al., 2003), whereby the finely detailed sequence characteristics of a dye-stained PCR product are revealed in the shape of a melting curve by post-PCR melting analysis (Rapley & Harbron, 2005). The principle of HRM is similar to that of real time PCR. However, the dye used in real time PCR, SYBR Green I, has some limitations, such as inhibiting the PCR and degrading the PCR products when SYBR Green I is at high concentration, and binding PCR amplicons unequally (Monis et al., 2005). Compared to SYBR Green I, the HRM fluorescent dyes, such as EvaGreen and SYTO9, do not inhibit or negatively affect PCR at higher concentrations, resulting in a more robust PCR signal. They are also able to bind GC-rich or AT-rich amplicons without sequence preference (Monis et al., 2005; Mao et al., 2007).

During HRM, the PCR products are gradually denatured by the following post-PCR melting program, commonly in 0.2°C increments, and the emitted fluorescence is measured by the instrument to generate a melting curve (Figure 4.1). Figure 4.1A shows that when the melting program reaches the dissociation temperature, the DNA strands are denatured and the fluorescent dye is released until the DNA is completely dissociated. Based on the fluorescence data (Figure 4.1A), the melt curve is generated to present the rate of change in fluorescence signal. If the PCR product is specific, the range of dissociation temperature is narrow and then the melt curve displays a sharp peak (Figure 4.1B), so the shape of the peak is an indication of the quality of the PCR product. The melting curves also reflect the characteristics of the PCR products, such as length, GC content, sequence, and heterozygosity (Rapley & Harbron, 2005). The melting curves can provide valuable information for mutation screening and genotyping (Wittwer et al., 2003). Such benefits make this method suitable for high-throughput genotyping and mutation screening (Dong et al., 2009b).

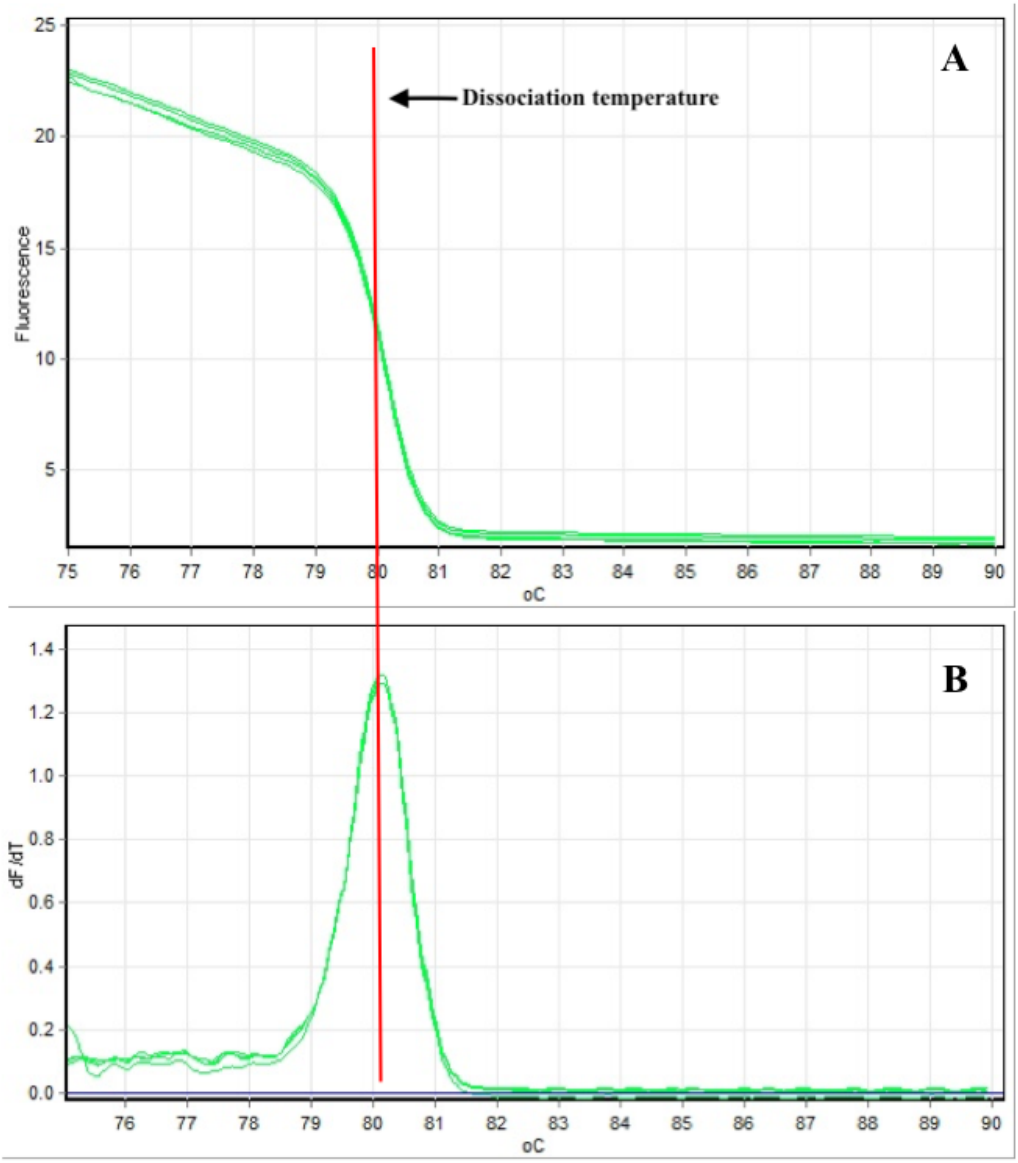


Figure 4.1 Raw channel (A) and melt curve (B) of HRM analysis. The dissociation temperature is indicated by the arrow. The melt curve reflects the rate of change in the fluorescence signal, so the shape of the peak in the melt curve can be used to judge the quality of the PCR product.

Perennial ryegrass is a wind-pollinated and self-incompatible species (Cornish 1979). Since each cultivar is derived as a synthetic population from different numbers of inbred lines, a perennial ryegrass cultivar is a heterogeneous population of genotypes with different degrees of heterozygosity (Momotaz et al., 2004). The aim of this project was to identify useful mutations in the target genes of interest from an EcoTILLING population including many cultivars. Because of the high degree of genetic diversity in perennial ryegrass, many individual plants in different New Zealand cultivars may have similar genetic backgrounds.

Hence, if the individuals in the EcoTILLING population could be pre-screened to maximize the number of different genotypes, the efficiency of the mutation screening can be increased. HRM analysis has been applied to genotyping with SSR markers in citrus (Distefano et al., 2013) and origanum (Mader et al., 2008).

However, HRM requires high specificity primers and short length PCR amplicons, up to 350 bp (Garritano et al., 2009). The amplicon length may affect the sensitivity and specificity of the subsequent post-PCR high resolution melting analysis, as it is harder to tell the relative change in fluorescence intensity for longer amplicons (Simko, 2016). Consequently, smaller amplicons are easier to analyse. Also, the high sensitivity of HRM analysis needs high specificity PCR amplification, because background noise reduces HRM resolution (Simko, 2016).

Targeted Induced Loci Lesions IN Genomes (TILLING) was developed to identify and analyse gene mutations in plants for functional studies. This technique was first developed in arabidopsis, and then applied to a range of crop plants (McCallum et al., 2000). Ethyl methane sulfonate (EMS) mutagenises can generate point mutations (single nucleotide polymorphism, SNP) in target genes of interest. EMS mutagenesis can produce multiple alleles within each gene, including missense, nonsense and splicing mutations (Liu et al., 2010).

EcoTILLING is the application of the TILLING technique to survey natural variation within ecotypes (Luca Comai et al., 2004). In this thesis, mutations within target genes will be detected in an EcoTILLING population. Once an individual with a useful allele is detected, this can be passed to a breeder, who can begin the process of backcrossing to develop a non-shattering phenotype. The individuals from the F2 generation, derived from the selected individual and another ryegrass plant, should contain the useful allele. These can then be backcrossed with the F1 generation to breed the F3 generation for homozygous useful alleles, and the non-shattering trait.

For TILLING (or Eco-TILLING), the critical step is to locate the mutation. In the TILLING protocol, the heteroduplex DNA, in which the two strands of DNA are derived from different DNA molecules, are cleaved at the mutation point by a purified endonuclease from celery, called CEL I, and are then detected by a LiCOR analyser or capillary electrophoresis platform (Till et al., 2006). In the LiCOR analysis, the forward and reverse PCR primers

labelled with a fluorescent dye are detected at 700 nm and 800 nm, respectively (Till et al., 2006). If PCR amplicons (up to 1.5 kb) with a point mutation are cut by CEL I nuclease, the cleavage amplicons can be detected, and the amplicons with a nucleotide change can be sequenced. Even though the TILLING approach has been applied with high-throughput in a large-scale population, the LiCOR analysis is time-consuming, usually over four hours per run, and the equipment and reagent costs need to be considered (Till et al., 2006).

In this chapter, firstly, a new method combining CEL I nuclease and HRM analysis was optimised for mutation screening of the target genes of interest. This method applied HRM analysis to check CEL I cleavage amplicons, rather than using LiCOR analysis or direct HRM analysis of PCR amplicons from EcoTILLING populations. Secondly, the full length of *LpCKX1*, which plays a key role in determining seed size, was cloned and the newly developed CEL I+HRM method was tested for mutation screening in *LpCKX1* within a ryegrass cultivar. Finally, SSR markers were used for HRM analysis to genotype within a perennial ryegrass cultivar. As the analyses indicated that genotyping was not going to be useful in the thesis, the data and analyses have been placed in Appendix 4.1.

4.2 Material and methods

4.2.1 Development of a method combining HRM analysis with CEL I nuclease to screen for mutations

4.2.1.1 CEL I nuclease isolation

Because the cost of the CEL I enzyme (Endonuclease from celery) is extremely high, about 10NZD per reaction for Guide-it Mutation Detection Kit (Cat. 631443), CEL I was extracted from 0.5 kg of celery bought at a supermarket and washed in cool water, as described by Till et al. (2006). All steps were performed at 4°C. After removing the white bottom and leaves, about 400 ml of celery juice was collected using a pre-chilled juice extractor to process pre-chilled celery sticks. The juice was centrifuged at 2600*g* for 20 min to pellet debris, and the supernatant then transferred to new 50 ml tubes. 1 M Tris-HCl and 0.1 M PMSF (phenylmethane sulfonyl fluoride) was added to the celery juice to make a final concentration of the solution with 0.1 M Tris-HCl, 100 μM PMSF, and pH 7.7. $(NH_4)_2SO_4$ powder was slowly added to the solution with gently stirring for 30 min, to a final concentration of 25%

saturation. The supernatant was then spun at 15000g at 4°C for 40 min, and the pellet discarded. The $(NH_4)_2SO_4$ concentration in the supernatant (25%) was adjusted to 80% saturation by slowly adding $(NH_4)_2SO_4$ powder and stirring gently for 30 min. The supernatant was spun at 15000*g* at 4°C for 1.5 h, and the pellet was then saved while being careful in decanting the supernatant. The pellet was resuspended in one tenth of the starting volume (about 40 ml) with Tris/KCl/PMSF buffer (100 ml of 1 M Tris-HCl, 100 ml of 5 M KCl, and 1 ml of 0.1 M PMSF, made up to a final volume of 1 L with water), and ensuring the pellet was completely dissolved. The resuspended solution was dialysed thoroughly against Tris/KCl/PMSF buffer for 1 h using a dialysis bag with a cutoff size of 12 kDa. After four replacements of the dialysis buffer, the enzyme solution was spun at 10000*g* for 20 min, and the supernatant with the CEL I was stored at -80°C. The extracted CEL I was checked on an SDS protein gel with Novex Sharp Pre-Stained Protein Standard (Life Technologies, USA).

The CEL I 10x digestion buffer was made based on Till et al. (2006). 50 ml of 10x digestion buffer contained 10 µl of 10 mg ml^{-1} BSA, 100 µl of 10% Triton X-100 (v/v), 2.5 ml of 2 M KCl, 5 ml of 1 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, pH7.5), 5 ml of 1 M $MgSO_4$ and 37.5 ml Milli-Q H_2O.

4.2.1.2 Optimisation of CEL I digestion

A pair of G/C plasmids supplied by Professor Song's lab in Yantai University, China, were used to optimise CEL I enzyme digestion conditions. The plasmids differ by one nucleotide at the same site: one plasmid has a guanine (G), and the other has a cytosine (C), so amplification of the gene provides known heteroduplexes to use to determine the efficacy of the purified CEL I. The plasmid sequences are attached in Appendix 4.2. Primers were designed with different amplicon lengths and containing the G/C site to optimise CEL I enzyme digestion conditions (Table 4.1).

All primers were checked by using Bioline Taq DNA polymerase with mixed G/C plasmids as template. The 15 µl of PCR master mix contained 1 Unit of Taq DNA polymerase (Bioline, UK), 1.5 µl Bioline 10x PCR buffer, 0.2 mM of each dNTP (Bioline, UK), 0.66 µM of each primer and about 50 pg of template plasmid. The reaction was subjected to 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 56°C annealing for 30 s, and 72°C extension for 80 s, and final 72°C extension for 10 min. The amplified products were separated on 1%

(w/v) agarose gel with TAE buffer at 90 V for 30 min. The G/C plasmid primers GC-CELI-F1/R1and GC-CELI-F/R373 were selected to optimise CEL I digestion conditions using different Bioline and Roche Taq DNA polymerases and the Type-it HRM kit.

Since CEL I only cuts at the mismatched point of heteroduplexes, the PCR products need to be hybridised. The DNA hybridisation program was 95°C for 5 min, then the temperature was initially decreased from 95°C to 85 °C at the rate of 2°C/sec and then reduced to 25°C at the rate of 0.1°C/sec. The digestion reaction contained 5 µl of hybridised PCR products, 2 µl of CEL I 10x digestion buffer, a variety of volumes of CEL I, with the final volume made up to 20 µl with Milli-Q H_2O. The reaction mixture was then incubated at 45°C with different digestion times, as shown in Table 4.2 and then stopped by adding 5 µl of 0.225 M EDTA. The digestion products were checked on a 2% (w/v) agarose gel with TAE buffer at 90 V for 30 min. Additionally, a commercial kit, Guide-it mutation Detection Kit (TaKaRa, Japan), was used as a control for the homemade CEL I nuclease. The details are attached in Appendix 4.3.

Table 4.1 G/C plasmid primer sequences for CEL I digestion optimisation

Primer pairs	**Amplicon length (bp)**	**Length (bp) to GC site**
GC-CELI-F1/R1	1465	829-GC-635
GC-CELI-F2/R1	1176	540-GC-635
GC-CELI-F3/R2	814	319-GC-494
GC-CELI-F4/R3	690	297-GC-392
GC-CELI-F/R373	373	156-GC-216
GC-CELI-F/R648	648	156-GC-491
GC-CELI-F/R902	902	156-GC-742
Name	Sequence	
GC-CELI-F1	ATCAAGACGATCTACCCGAGTAACAAT	
GC-CELI-F2	TGGCGAACAGTTCATACAGAGTCTT	
GC-CELI-F3	AGGAAGGTGGCTCCTACAAATGC	
GC-CELI-F4	CCATCATTGCGATAAAGGAAAGGCTAT	
GC-CELI-R1	CCTCCTTACCACCAGCGAATCC	
GC-CELI-R2	GCTGGAGAGGGAAGTGTGATGAC	
GC-CELI-R3	GTGAGTTGAGGATGAACCTGAGAAGTA	
GC-CELI-F	TATCTCCACTGACGTAAGGGATGACG	
GC-CELI-R373	CCGTTTCCCTCTCTTGGCG	
GC-CELI-R648	GGAGAGGGAAGTGTGATGACCCAAATGA	
GC-CELI-R902	TTAGAAACCATCTTGAAACCGATTGAT	

Table 4.2 The optimisation of CEL I digestion with different amounts of CEL I and reaction time.

	Bioline Taq	Roche Taq	Type-it HRM Kit
Primers	GC-CELI-F1/R1	GC-CELI-F1/R1	GC-CELI-F/R373
Digestion Time (min)	15, 30, 60, 90, 120	15, 30, 60, 90, 120	15, 30, 60, 90, 120
Amount of CEL I (µl)	0.2, 0.5, 1.0, 1.5, 2.0	0.2, 0.5, 1.0, 1.5, 2.0	0.2, 0.5, 1.0, 1.5, 2.0

4.2.1.3 Trials of HRM analysis to identify CEL I digestion products

4.2.1.3.1 Trial 1

The CEL I digestion products were checked on agarose gel and through HRM analysis. The PCR products from the GC-CEL I-F1/R1 primers were from two template DNAs: G/C mixed plasmid and only G plasmid by using Bioline Taq DNA polymerase. The PCR products were hybridised (Section 4.2.1.2). The CEL I digestion reaction contained 15 µl of hybridised PCR products, 1 µl of CEL I or H_2O (Table 4.3), 2 µl of digestion buffer and final volume made to 20 µl with Milli-Q H_2O. The digestion reaction was incubated at 45°C for 30 min and stopped by addition of 5 µl of 0.225 M EDTA. The digested products were checked on a 2% (w/v) agarose gel with TAE buffer at 10 V/cm for 30 min. The digestion products were subjected to HRM analysis with 1.5 µl of EvaGreen (20x in water) in each reaction. The EvaGreen dye and SYBR Green dye were compared and then the EvaGreen dye was chosen as the fluorescent dye to detect CEL I digestion products by using HRM analysis. More details are shown in Appendix 4.4.

Table 4.3 CEL I digestion with different treatment of two PCR products.

	Treatment			
Components	1	2	3	4
Template DNA	G/C	G	G/C	G
CEL I (µl)	1	1	0	0

4.2.1.3.2 Trial 2

The PCR products from primer pairs GC-CELI-F/R373 and F/R648 were amplified by using Roche Taq DNA polymerase (Roche, Germany). Each 15 µl PCR reaction contained 0.6 Unit of Taq DNA polymerase, 1.5 µl Roche 10x PCR buffer, 3 mM of $MgCl_2$ (Bioline, UK), 0.2

mM of each dNTP (Bioline, UK), 0.66 µM of each primer and about 50 pg of template plasmid. The PCR products were hybridised (Section 4.2.1.2)

Two hybridised PCR products were digested with different treatments, as shown in Table 4.4. For Treatments 1 and 2, 1 µl of 20x EvaGreen dye was added just before HRM analysis, but for Treatment 3, 1 µl of 20x EvaGreen was added into the digestion reaction mixture before incubation. All digestion reactions were incubated at 45°C for 30 min and then stopped by adding 5 µl of 0.225 M EDTA. The digestion products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 45 min, and by HRM analysis. The HRM analysis was repeated with different increments of temperature including 0.1°C/2s, 0.1°C/6s, and 0.1°C/10s.

Table 4.4 Assessments of CEL I digestion using PCR products from G/C plasmid.

Components (µl)	**Treatment**			**Control**
	1	2	3	
PCR products	15	5	5	5
10x CEL I buffer	2	2	2	2
CEL I	0.5	0.5	0.5	0
EvaGreen dye	0 [a]	0 [a]	1 [b]	0
Final volume with H_2O	20	20	20	20

[a] 1 µl of 20x EvaGreen dye was added just before HRM analysis
[b] 1 µl of 20x EvaGreen dye was added before CEL I digestion reaction

4.2.1.3.3 Trial 3: Confirmation of CEL I digestion products using HRM analysis

The PCR was performed with GC-CELI-F/R373 and the three template plasmids, G/C, G/G, and C/C, and using Roche Taq DNA polymerase. Each reaction was 15 µl containing 0.6 Unit of Roche Taq DNA polymerase, 1.5 µl Roche 10x PCR buffer, 3 mM of $MgCl_2$, 0.2 mM of each dNTP (Bioline, UK), 0.66 µM of each primer and about 50 pg of template plasmid. The reaction was subjected to 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 56°C annealing for 30 s, and 72°C extension for 80 s, and final 72°C extension for 10 min. PCR products were hybridised (Section 4.2.1.2). Ten hybridised PCR products were digested with different treatments, as shown in Table 4.5. The CEL I digestion reaction was incubated at 45°C for 30 min, then stopped by adding 5 µl of 0.225 M EDTA. 10 µl of digestion products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 30 min. 0.8 µl of

20x EvaGreen were added into 15 µl of digestion products and then HRM were performed from 75-90°C with three different temperature increments 0.1°C/2s, 0.1°C/6s, and 0.1°C/10s.

Table 4.5 Treatments of 10 CEL I digestion reactions

No.	Digestion Template	Template volume (µl)	CEL I (µl)	10x digestion buffer (µl)	H_2O (µl)	Final volume (µl)
1	G/C	15	0.5	2	2.5	20
2	G/C		0.5		2.5	
3	G/C		0.5		2.5	
4	G/C		0		3	
5	G		0.5		2.5	
6	G		0.5		2.5	
7	C		0.5		2.5	
8	C		0.5		2.5	
9	G		0		3	
10	C		0		3	

4.2.1.3.4 Trial 4: CEL I digestion using different ratios of heterozygous alleles

The PCR was performed with primer GC-CELI-F/R373 and plasmid DNA, either G or C, and using Roche Taq DNA polymerase. The mixed template PCR products were hybridised (Section 4.2.1.2) to form heteroduplexes with a mismatched nucleotide which can be digested by CEL I nuclease. The template mixture was digested by CEL I and each digestion reaction contained 2 µl of 10x digestion buffer, 1 µl of CEL I, 15 µl of the mixed template, and 2 µl of Milli-Q H_2O (Table 4.6). The digestion reaction was incubated at 45°C for 30 min and stopped by addition of 5 µl of 0.225 M EDTA. The digestion products were checked by a 3% (w/v) agarose gel at 100 V for 30 min. 15 µl of digestion products were added to 0.8 µl of 20x EvaGreen and HRM performed from 75-90°C.

Table 4.6 Ratios of G and C plasmid PCR products in the CEL I digestion reaction

Ratios of G/C PCR product	Volume of G (µl)	Volume of C (µl)
0:10	0	15
1:9	1.5	13.5
2:8	3	12
3:7	4.5	10.5
4:6	6	9
5:5	7.5	7.5

4.2.2 Testing the CEL I+HRM method using *LpCKX1*

4.2.2.1 Cloning of *LpCKX1* and *LpCKX2* genes

Sequences of candidate *CKX1* and *CKX2* from perennial ryegrass were determined through Nucleotide BLAST searching a ryegrass transcriptome database (Section 3.2.3). The putative *LpCKX1* and *LpCKX2* sequences were aligned with their subfamily members in other Poaceae and the Neighbour-Joining phylogenetic trees were constructed by using Clustal X (1.83) software.

The full cDNA sequence of *LpCKX1* and the second half cDNA sequence of *LpCKX2* were identified from the perennial ryegrass transcriptome database. Based on the identified cDNA sequences, the primers for amplifying the exon and intron parts of both genes were designed and optimised (Table 4.7). The full length of *LpCKX1* and part of *LpCKX2* were amplified using genomic DNA from three individuals of cv. Nui using Platinum Taq DNA polymerase (Invitrogen, USA) and primer pairs, LpCKX1-F1R1 and LpCKX2-F2R3. The PCR products were then ligated with a TOPO TA Cloning Kit (Invitrogen, USA) following the instruction manual. The constructed plasmids were transformed into *E. Coli* DH5α competent cells. The plasmids were isolated from selected transformed cells using DNA-spin Plasmid DNA Purification kit (iNtRON Biotechnology Inc. South Korea) following the manual. All plasmids were sent to Macrogen for sequencing and the results were aligned using Clustal X. Gene structures of *LpCKX1* and *LpCKX2* were determined.

Table 4.7 Primer sequences for *LpCKX1* and *LpCKX2*

Gene	Primer name	Sequences	Amplicon region
LpCKX1	LpCKX1F1	GCCAGCCTTATCGCGCCTAA	Full length
	LpCKX1F2	AATTACCAACACATCGTCACGTTTCTC	
	LpCKX1R1	CAGTCTAGTTGAAGATGTCCTGTCCT	
	LpCKX1R2	GTCCCACTTGCCGCCGAAAT	
	LpCKX1F3	GGCGAGCAGATGTGGATCGA	Intron 1
	LpCKX1F4	ATGTGGATCGACGTGCTGCG	
	LpCKX1R5	AGGTGGCGAAGTCGGTGTAGA	
	LpCKX1R6	GAAGGTGGCGAAGTCGGTGTA	
	LpCKX1F5	CCCGTGGCTCAACTTGCTCGTG	Intron 2
	LpCKX1F6	TCCCGCATCGCCGACTTCGAC	
	LpCKX1R3	TGCGCTTGCAGCCGCTTCAG	
	LpCKX1R4	GTCGCAGAAGCGCAGGATCTTCTG	
	LpCKX1F7	GTGTTACTCGTCGCGCTAATCG	Exon 1
	LpCKX1F8	ACGTGTTACTCGTCGCGCTAAT	
	LpCKX1F11	CGTGTTACTCGTCGCGCTAATCG	
	LpCKX1R7	CGGAATGTCTGGCCGCTGAT	
	LpCKX1R8	AATGTCTGGCCGCTGATCCC	
	LpCKX1R11	GGCCGTGCCGGAATGTCTGG	
	LpCKX1R12	GACGGTGAGGTGGAGGTAGTCC	
	LpCKX1R13	GGTGAGGTGGAGGTAGTCCGTCC	
	LpCKX1F9	ACGGTGACGTGCTCCAAGTCC	Exon 2
	LpCKX1F10	AGACGGTGACGTGCTCCAAGT	
	LpCKX1R9	ATTTGTTGAGCGGGTAGATGACCATAG	
	LpCKX1R10	GGATTTGTTGAGCGGGTAGATGAC	
	LpCKX1F12	CTTCGGCGGCATCCTCCAG	Intron 2 to exon 3
	LpCKX1F13	CGTCTTCGGCGGCATCCTC	
	LpCKX1R14	AGTCTAGTTGAAGATGTCCTGTCCTG	
	LpCKX1R15	TCAGTCTAGTTGAAGATGTCCTGTCC	
LpCKX2	LpCKX2F1	GATGTGCTGCTGCGGGAGTT	Exon 3
	LpCKX2R1	GCTATACGTCGGCATCGATCAATC	
	LpCKX2R2	CATGAGAATCCCTCGGCGTCTAG	
	LpCKX2F2	CGCTCGTCCTTCTTCTCGGA	Exon 2 to 3' UTR
	LpCKX2F3	TCGTCCTTCTTCTCGGAGGC	
	LpCKX2R3	ACATCGTTTCCCTTGCCATGATAT	
	LpCKX2R4	TTTAACATCGTTTCCCTTGCCATG	
	LpCKX2F4	AGCTGGCGCTCGTCCTTCTT	Intron 2
	LpCKX2F5	TGGCGCTCGTCCTTCTTCTC	
	LpCKX2R5	ATGCGGCACTTCCCACAAACC	
	LpCKX2R6	GGATGCGGCACTTCCCACAAA	

The region of exon 2 and the region from intron 2 to exon 3 in *LpCKX1* were sequenced from cv. Nui. The primers were designed and optimised (Table 4.7). The PCR for the two regions

of *LpCKX1* was performed using Platinum Taq DNA polymerase (Invitrogen, USA), with genomic DNA from 11 individuals of cv. Nui and primer LpCKX1-F10/R10 and LpCKX1-F13/R15. The PCR products were separated on a SYBR Safe stained 1% (w/v) agarose gel with TAE buffer, at 7 V/cm for 40 min, and were purified using UltraClean 15 DNA Purification Kit (Mo Bio Laboratories, USA) following the instruction manual. All samples were sequenced by Macrogen.

4.2.2.2 HRM detection of CEL I digestion products of *LpCKX1*

Exon 1 of *LpCKX1* was selected to test the mutation screening method with CEL I + HRM, because the sequencing results from exon 2 of *LpCKX1* showed many SNPs and exon 3 was not long enough for this experiment. The nested PCR approach was applied, and the first PCR run was carried out using Platinum DNA polymerase, with genomic DNA from 20 individual plants from of cv. Nui and primer LpCKX1-F1R1. Each reaction contained 0.2 Unit of Platinum Taq DNA polymerase, 1 µl Platinum 10x PCR buffer, 1.5 mM of $MgCl_2$, 5% of KB Extender, 0.2 mM of each dNTP (Bioline, UK), 0.5 µM of each primer and about 200 ng of template DNA in a total volume of 10 µl. The reaction was subjected to 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 62°C annealing for 30 s, and 72°C extension for 3 min, and final 72°C extension for 6 min. The amplified products were diluted 100 times with Milli-Q water for the following mutation screening.

The mutation screening method was tested in two trials (Table 4.8). Both trials used Roche Taq DNA polymerase and primer LpCKX1-F11R11. In Trial 1, the reaction was carried out with standard PCR and without EvaGreen Dye, but Trial 2 was performed by HRM analysis with EvaGreen dye. In addition, the two trials used different DNA templates. In Trial 1, the template DNA in each reaction was a mixture of PCR products from the first PCR run: PCR products from Nui 1 to Nui 20 were added to PCR product from Nui 1, which was set as control, in the ratio of 1 to 1. Each reaction of Trial 1 contained from one to four alleles. In Trial 2, the template DNA for each reaction was only from the diluted first run PCR products, so each reaction contained one or two alleles. For Trial 1, the reaction was subjected to 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 62°C annealing for 30 s, and 72°C extension for 40 s, and final 72°C extension for 6 min. For Trial 2, the reaction was carried

out using Rotor-Gene Q with the same PCR program as in Trial 1, followed by the HRM analysis program from 75-95°C.

The PCR products was subjected to DNA hybridisation to produce heteroduplexes for CEL I cleavage (Section 4.2.1.2). The digestion reaction contained 15 µl of hybridised PCR products, 2 µl of CEL I 10x digestion buffer, 1 µl of CEL I nuclease, and made up to a final volume of 20 µl with Milli-Q H_2O. The reaction mixture was then incubated at 45°C for 30 min and then stopped by addition of 5 µl of 0.225 M EDTA. The 8 µl of digestion products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 50 min. The remaining digestion mixture was subjected to HRM analysis from 80-99°C.

Table 4.8 Components of two trails for detection of CEL I+HRM method

	Trial 1	**Trial 2**
Components	Volume per 15 µl reaction (µl)	
Roche Taq DNA polymerase	0.12	0.12
Roche 10x PCR buffer	1.5	1.5
50 mM $MgCl_2$	0.45	0.9
20 mM dNTP	0.15	0.15
KB Extender	0.75	0.75
20X EvaGreen	0	0.75
Primer	1	1
Template DNA	1[a]	1[b]
PCR-grade water	15	15

[a] 20 PCR products from the first run of nested PCR were mixed with the PCR product from Nui 1, with the same ratio of volume.
[b] the template DNA for each reaction was only from first run PCR products, not mixed.

4.3 Results

4.3.1 HRM analysis in combination with CEL I

The drawbacks of TILLING have been discussed in Section 4.1, but the particularly time-consuming step is the detection of the cleavage products cut by CEL I nuclease. The HRM analysis can detect differences in DNA fragments, so the aim of this section was to apply HRM analysis to detect the cleavage products cut by CEL I, instead of using a gel electrophoresis platform.

4.3.1.1 CEL I enzyme isolation and optimisation

The SDS protein gel for the isolated CEL I is shown in Figure 4.2. Two bands can be observed in the CEL I lane, at about 48 kDa and 52 kDa, compared with Novex Sharp Pre-Stained Protein Standard ladder. The primers for G/C plasmids were optimised (Figure 4.3). The primers, GC-CELI-F1/R1 and GC-CELI-F/R373, were selected for the optimisation of CEL I digestion conditions with different Taq PCR products, because the lengths of two PCR products showed the largest difference.

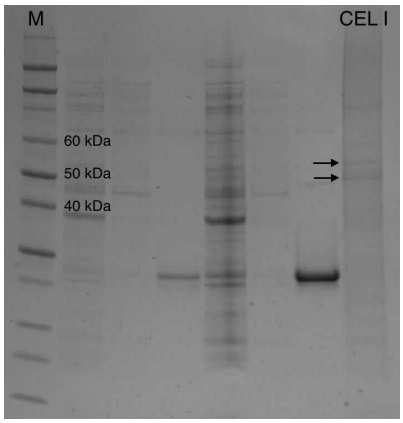


Figure 4.2 SDS protein gel of CEL I. Two bands of CEL I are indicated by black arrows at about 48 kDa and 52 kDa, respectively (right lane). M: Novex Sharp Pre-Stained Protein Standard ladder (left lane).

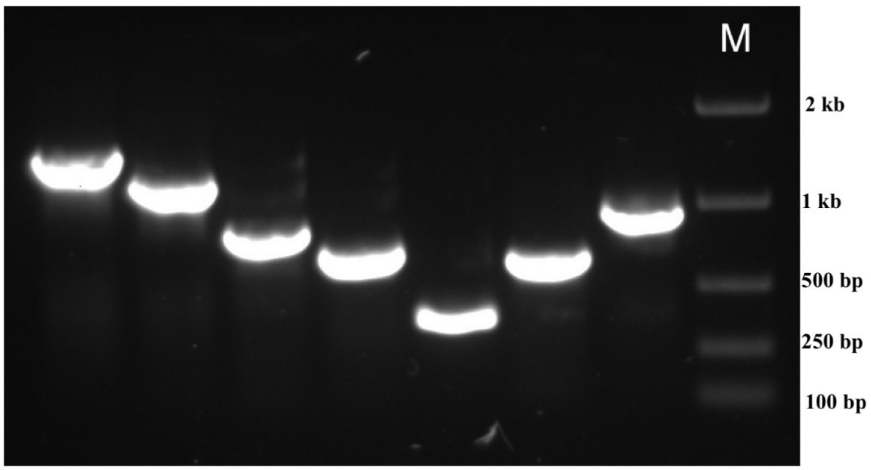


Figure 4.3 The PCR products of GC-CELI-F1/R1, F2/R1, F3/R2, F4/R3, F/R373, F/R648, F/R902 amplified with G/C plasmids from left to right, respectively (refer to Table 4.1 for sequences). All PCR products were checked on a 1% (w/v) agarose gel with TAE buffer at 10 V/cm. M: EasyLadder I.

The PCR products were digested by CEL I with different amounts of enzyme, different Taq polymerases and incubation time. The digested products were checked by gel electrophoresis. From the three gel photos shown in Figure 4.4, the cleavage products indicated by arrows can be observed clearly, but the same digestion products show smeared bands. For Bioline Taq PCR product (Figure 4.4A), the digested products show two bright bands at the correct position, and the optimal digestion condition could be 0.5 µl of CEL I for 30 min. Interestingly, no matter how much CEL I was added, the digestion results were not affected by digestion times greater than 30 min. The digestion result from Roche Taq is shown in Figure 4.4B, and two cleavage products can be observed, but almost all digestion products show smeared bands when the more than 0.2 µl CEL I was added. The optimal digestion condition for Roche Taq could be 0.2-0.5 µl of CEL I for 30 min.

For the Type-it HRM kit (Figure 4.4C), even though the two digested bands can be observed, the bands are very weak, compared with the other two Taq polymerases. The reason behind the weak bands in Figure 4.4C maybe because after 30 min of gel electrophoresis the digestion products of GC-CELI-F/R373, are no longer visible. Compared with other digestion products in Figure 4.4C, the best digestion condition for the Type-it HRM Kit could be 0.5 µl of CEL I for 30 min, as the cleavage products in the lane can be observed clearly. Even though, the brightest cleavage products are from the lane with products digested for 15 min with 1µl CEL I, but the digestion template in the same lane is also brighter than other templates. The optimisation of three different master mixes show different digestion results, suggesting that the components of the digestion reaction significantly affect the CEL I digestion results.

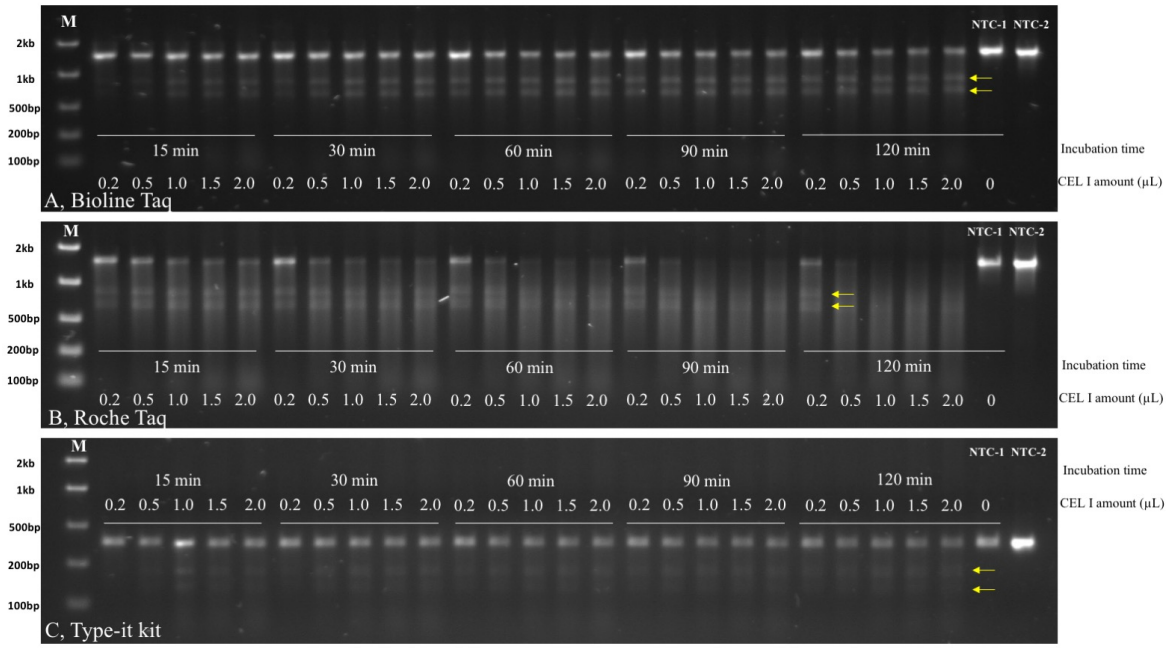


Figure 4.4 The cleavage products from three different master mixes. A: PCR products of GC-CELI-F1/R1 from Bioline Taq homemade master mix; B: PCR products of GC-CELI-F1/R1 from Roche Taq homemade master mix; C: PCR products of GC-CELI-F/R373 from the Type-it HRM kit. NTC-1: no CEL I added into the reaction; NTC-2: PCR products. All PCR products were checked on a 2% (w/v) agarose gel with TAE buffer at 10 V/cm for 30 min. M: EasyLadder I. Cleavage products are indicated by arrows.

4.3.1.2 Optimisation of HRM analysis of CEL I digestion products

4.3.1.2.1 Trial 1: HRM analysis for CEL I digestion product detection

As HRM analysis is a highly sensitive method used to detect DNA fragment size by dissociation temperature, the components of a reaction mixture could affect the HRM result. The PCR products from primer GC-CELI-F1/R1 with different template plasmids were digested by CEL I (Table 4.3). Digestion products were checked by agarose gel electrophoresis and HRM analysis. From the gel photo (Figure 4.5), the G/C plasmid product has been cleaved, and the two digestion bands are clearly shown, whereas the G plasmid products have no digested products, but lane 2 also has a smear, suggesting that CEL I could have degraded the double strand of DNA. In Figure 4.5 lanes 3 and 4, there are no cleavage products or smeared bands visible. From the melting curve of HRM analysis (Figure 4.6), the dissociation temperature of the DNA fragments significantly increased with the addition of CEL I digestion buffer alone, or the addition of both CEL I enzyme and digestion buffer. Multiple peaks were shown on melting curves for each sample, but no peaks can be identified

as cleavage products cut by CEL I. In addition, the PCR products from GC-CELI-F1/R1 without CEL I treatment also show multiple peaks, which means non-specific products were amplified from the primer pair.

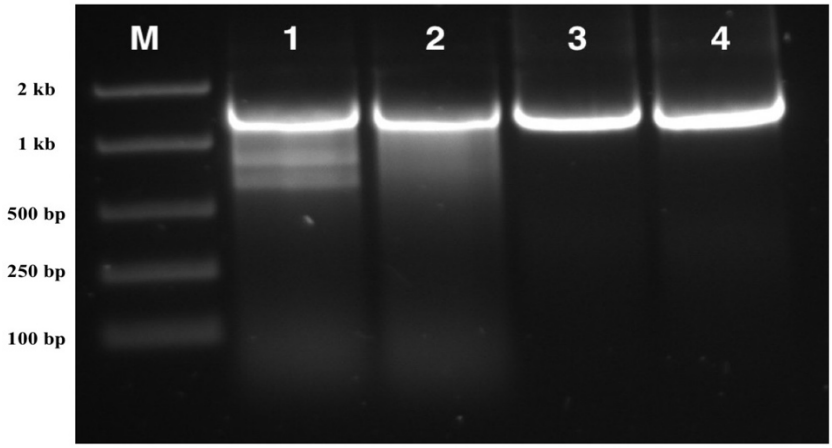


Figure 4.5 CEL I digestion products with different treatment refer to Table 4.3. 1: G/C plasmids with 1 µl of CEL I; 2: G plasmid with 1 µl of CEL I; 3: G/C plasmids without CEL I, but with digestion buffer; 4: G plasmid without CEL I, but with digestion buffer. All PCR products were checked on a 2% (w/v) agarose gel with TAE buffer at 10 V/cm for 30 min. M: EasyLadder I.

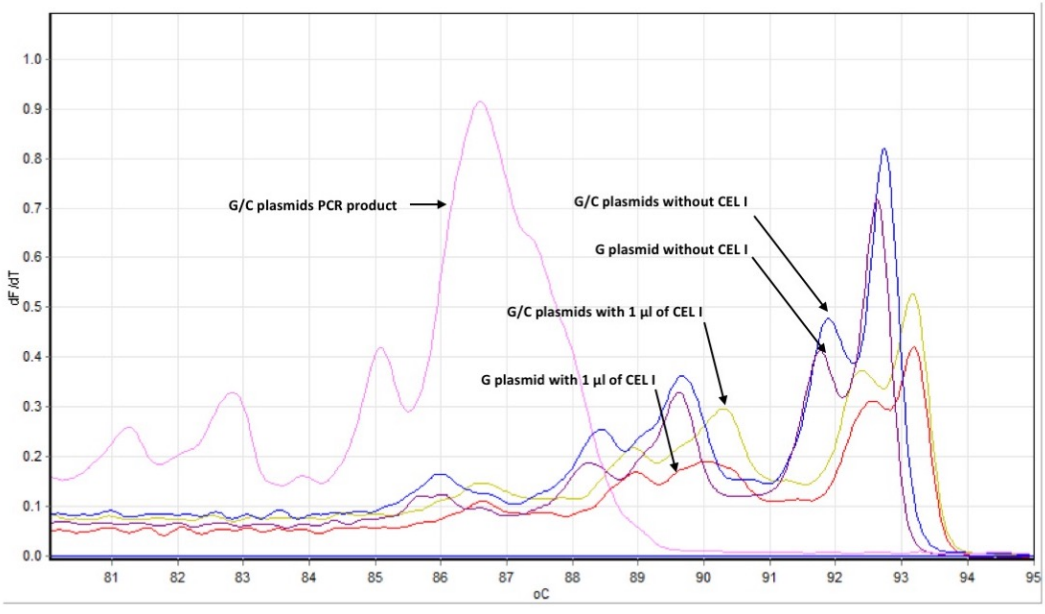


Figure 4.6 Melting curve in HRM analysis of CEL I digestion products with different treatments, refer to Table 4.3. Pink: G/C plasmids PCR product; Yellow: G/C plasmids with 1 µl of CEL I; Red: G plasmid with 1 µl of CEL I; Blue: G/C plasmids without CEL I, but with digestion buffer; Purple: G plasmid without CEL I, but with digestion buffer.

4.3.1.2.2 Trial 2

Two G/C plasmid primers, GC-CELI-F/R373 and F/R648, were used for PCR amplification and the products were digested by CEL I enzyme using different quantity of PCR products and EvaGreen dye (Table 4.4). The cleavage products from Treatment 1 are stronger than from Treatments 2 and 3 (Figure 4.7), which means that a greater amount of DNA in the digestion reaction could improve assessment by agarose gel electrophoresis. Compared with digestion results from Treatment 2, the digestion products in Treatment 3 is weaker and more smeared, which means that the EvaGreen fluorescent dye could inhibit CEL I digestion reaction.

From the HRM analysis, the melting results from primer FR373 are significantly better than from primer FR648, because of high specificity (Figure 4.8): the cleavage products from the three treatments were detected and are indicated by arrows. The peaks from Treatment 1 (red line in Figure 4.8) show higher fluorescent singles and higher denature temperature. The reaction of Treatment 1 contained different components compared with Treatments 2 and 3, as 15 µl of PCR products were added into the digestion reaction. Interestingly, Treatments 2 and 3 do not show too much difference in the melting curves (blue: Treatment 2, pink: Treatment 3), although fluorescence signal from Treatment 3 is slightly higher than Treatment 2. The HRM analysis of Treatments 2 and 3 is not compatible with the results from gel electrophoresis (Figure 4.7). Taken together, from all optimisation results of CEL I digestion, the use of a greater amount of digestion template should improve the digestion results, and EvaGreen fluorescent dye can be used to detect the cleavage products by HRM analysis.

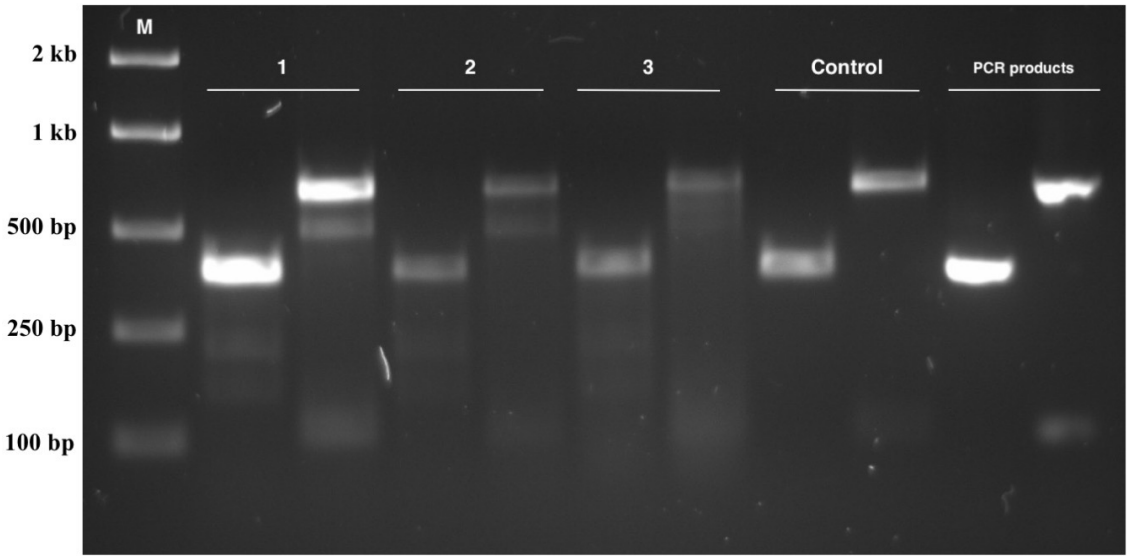


Figure 4.7 The PCR products from GC-CELI-F/R373 and F/R648 with different treatments in EvaGreen and template amount for CEL I digestion. 1: 15 µl of template for digestion and no EvaGreen added before digestion reaction; 2: 5 µl of template for digestion and no EvaGreen before digestion reaction; 3: 5 µl of template for digestion and 1 µl of 20x EvaGreen added before digestion reaction; Control: 5 µl of template for digestion, but no CEL I added. All PCR products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 45 min. M: EasyLadder I.

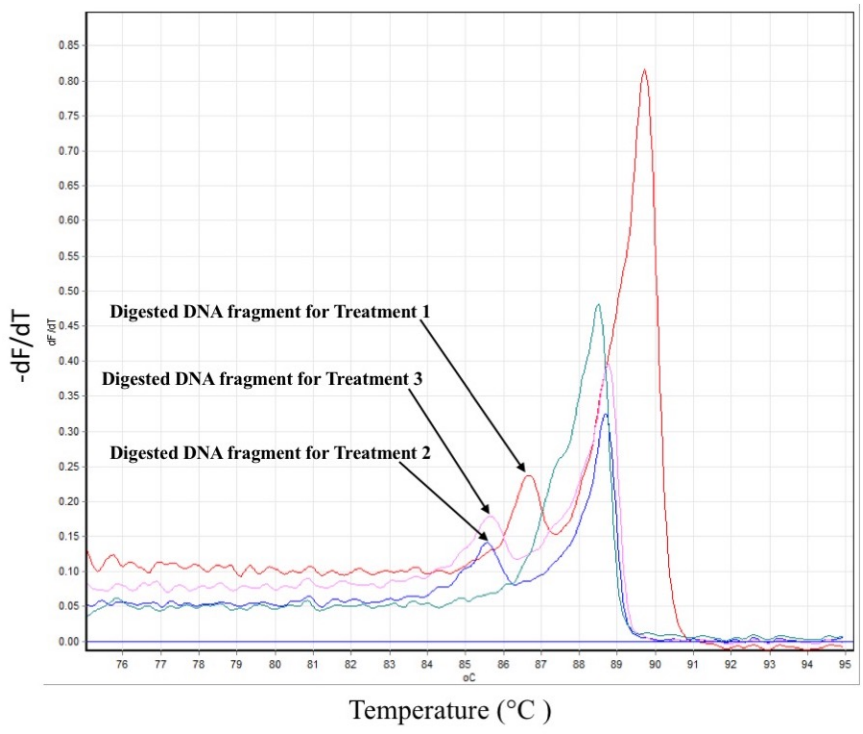


Figure 4.8 The melting curves of CEL I digestion of PCR products from GC-CELI-F/R373 with different Treatments. Red: Treatment 1; Blue: Treatment 2; Pink: Treatment 3; Green: control. For details of each Treatment refer to Table 4.4.

4.3.1.2.3 Trial 3: Confirmation of the method combining CEL I digestion with HRM analysis

PCR products from three different template plasmids, including G/C, G/G, and C/C plasmids, were digested by CEL I (Table 4.5). Two cleavage products from G/C digestion templates can be identified clearly, as indicated by arrows in Figure 4.9. The melting curves of the ten digestion products are presented in Figure 4.10. The cleavage products of G/C digestion templates can be identified in the melting curves, consistent with the results of gel electrophoresis. These results suggest that the innovative method combining HRM analysis with CEL I digestion can be applied to mutation screening.

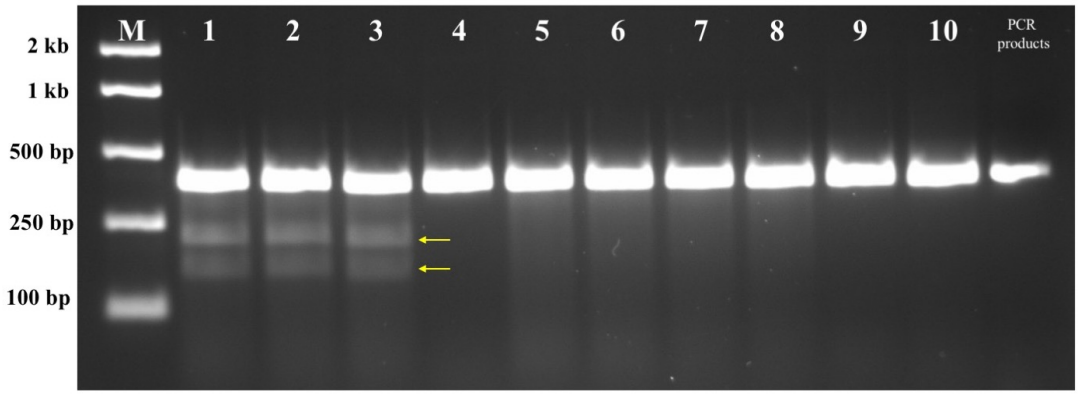


Figure 4.9 The ten PCR products with different treatments. 1-3: G/C plasmid products with 0.5 µl of CEL I, 4: G/C plasmid product without CEL I, 5-6: G plasmid products with 0.5 µl of CEL I, 7-8: C plasmid products with 0.5 µl of CEL I, 9-10: G and C plasmid products without CEL I. All PCR products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 30 min. M: EasyLadder I. Cleavage products are indicated by arrows.

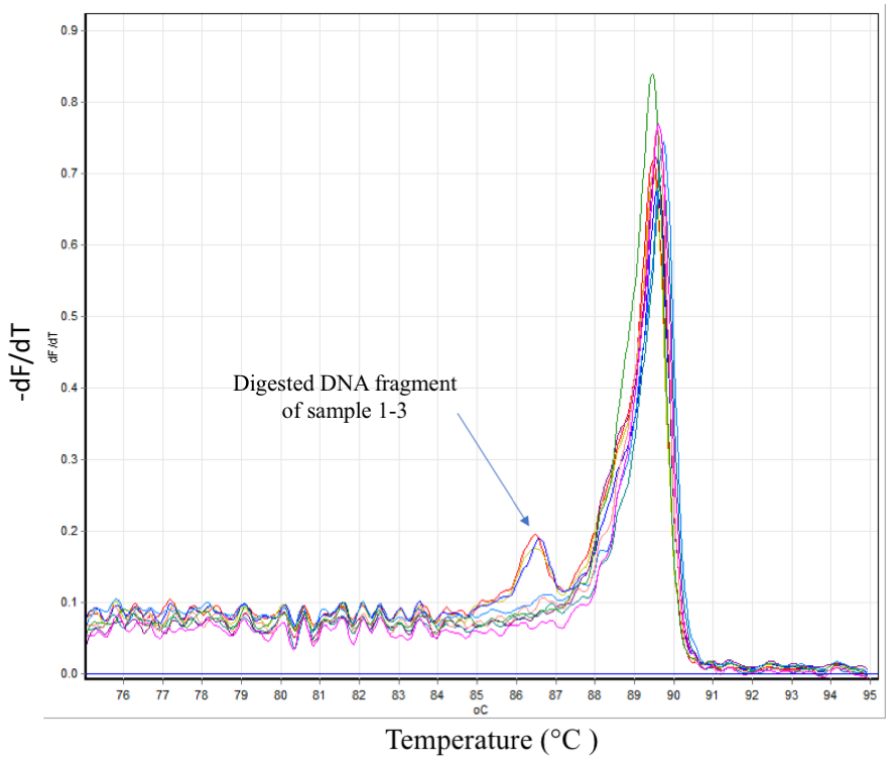


Figure 4.10 The melting curve of ten digestion reactions with different treatments as described in Table 4.5 and Figure 4.9. The digested band peaks are indicated by the arrow.

4.3.1.2.4 Trial 4: CEL I digestion with different ratios of heterozygous alleles

Different ratios of the G and C plasmid PCR products were mixed as digestion templates. When the ratio was close to 50:50, more heteroduplexes can anneal, resulting in the digestion products showed stronger bands (Figure 4.11). With the ratio of digestion templates (G:C) approaching 5:5, the peak of cleavage products (light blue in Figure 4.12) is higher in the melting curve analysis, consistent with the gel electrophoresis. According to these results, the range of G/C digestion template ratio from 2:8 to 5:5 could be detected by gel electrophoresis and HRM analysis.

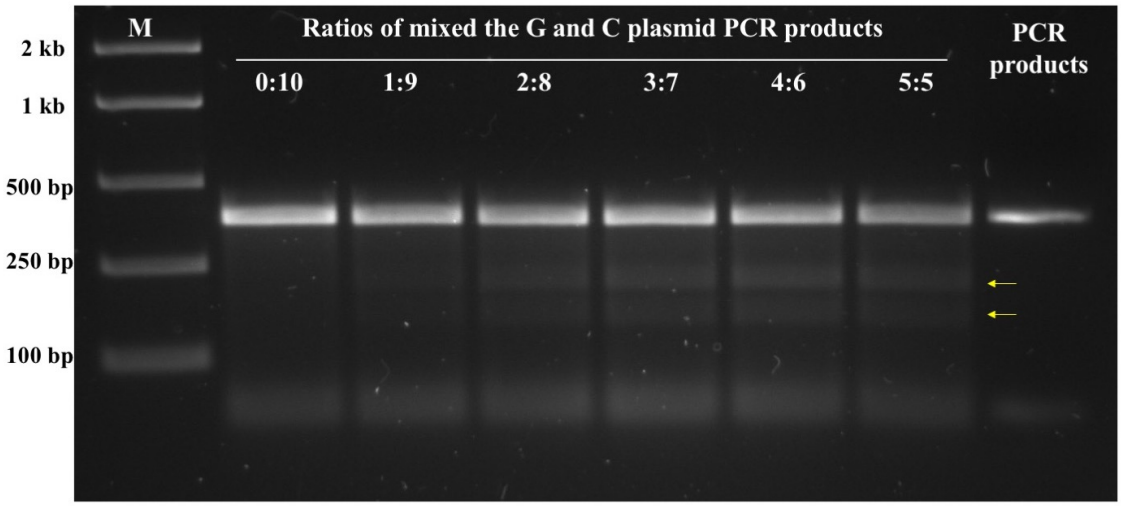


Figure 4.11 The digestion products from the different ratios of mixed PCR products. All PCR products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 30 min. M: EasyLadder I. Cleavage products are indicated by arrows.

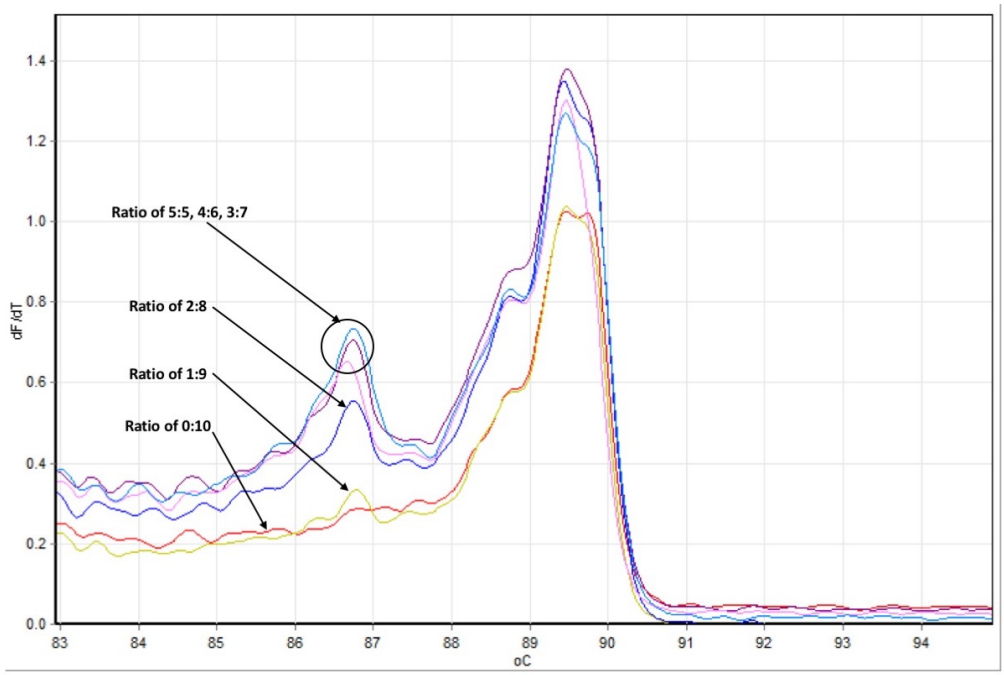


Figure 4.12 The melting curves of digestion samples from the different ratios of digestion template. Light blue: the digestion template ratio of 5:5; purple: the ratio of 4:6; pink: the ratio of 3:7; blue: the ratio of 2:8; yellow: the ratio of 1:9; red: the ratio of 0:10.

4.3.2 Sequencing of full length of *LpCKX1* and part of *LpCKX2*

The full length of *LpCKX1* and exon 2 to the 3' UTR of *LpCKX2* were sequenced. Their DNA sequences are in Appendices 4.5 and 4.6. According to the sequencing results, *LpCKX1* has a high GC-content of about 62%. The Neighbor-Joining phylogenetic trees of *CKX1* and *CKX2* were constructed and are shown in Figure 4.13. Multiple sequence alignment of both genes showed that *LpCKX1* and *LpCKX2* have close relationships with homologues from other Poaceae.

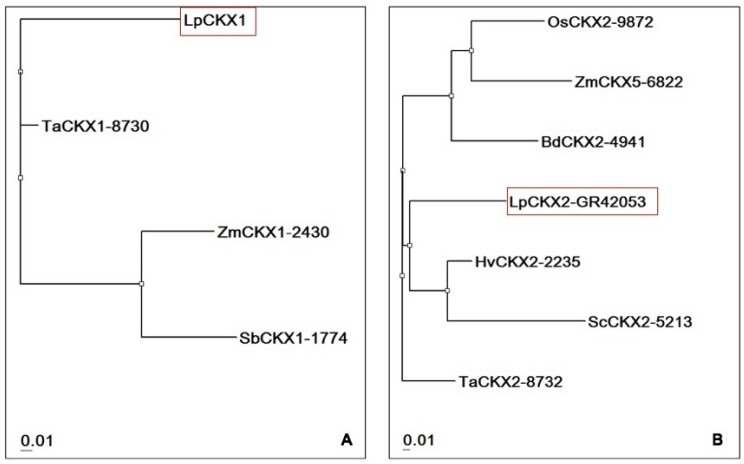


Figure 4.13 Phylogenetic tree of *CKX1* (A) and *CKX2* (B) sub-family in monocots. The *LpCKX1* and *LpCKX2* are boxed with red. *Ta, Triticum aestivum; Os, Oryza sativa; Lp, Lolium perenne; Hv, Hordeum vulgare; Sb, Sorghum bicolor; Zm, Zea mays; Si, Setaria italica; Bd, Brachypodium distachyon; Sc, Secale cereale.*

The gene structures of *LpCKX1* and *LpCKX2* are shown in Figure 4.14, and the location of designed primers (Table 4.7) are indicated at their approximate sites. Full length sequences of *LpCKX1* genes amplified from three genomic DNAs from cv. Nui were also obtained. The sequencing results of *LpCKX1* were aligned with the cDNA sequence from the perennial ryegrass transcriptome database (Appendix 4.7). Over 15 SNPs were identified in exon 1 and exon 2, respectively, while only two SNPs were detected in exon 3, suggesting that exon 3 may be highly conserved.

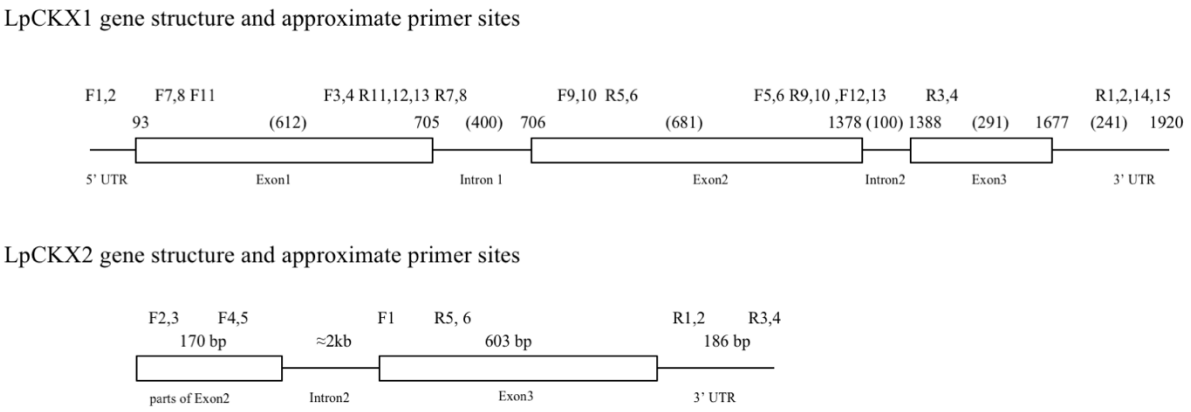


Figure 4.14 Gene structures of *LpCKX1* and *LpCKX2* and the approximate sites of the primers.

The PCR products from cv. Nui and primer LpCKX1-F10R10 were sequenced and aligned (Appendix 4.8). From the alignment results, 19 SNPs were identified in exon 2, suggesting that this region cannot be used for mutation screening. The PCR products from 11 individuals from cv. Nui and primer LpCKX1-F13R15 were sequenced and aligned (Appendix 4.9). From the alignment results, two SNPs and one INDEL were identified in exon 3. This suggests that the DNA sequence in the region of LpCKX1-F13R15 is conserved, especially compared with DNA sequences of exon 2 in *LpCKX1*.

Exons 1 and 2 of *LpCKX1* were suitable to test the mutation screening method with HRM analysis and CEL I nuclease (CEL I + HRM), because the lengths of both exons are suitable for CEL I digestion, being about 650 bp. However, as many mutations were detected in exon 2 of *LpCKX1*, exon 1 was selected to test the CEL I+HRM method.

4.3.3 Detection of CEL I digestion products from *LpCKX1*

Exon 1 of *LpCKX1* was used to test if the method of combined HRM analysis and CEL I digestion could detect the mutations. From Trial 1, many cleavage bands can be seen on the agarose gel (Figure 4.15A), which means that there were many mismatched nucleotides in the heteroduplexes. These mismatched nucleotides meant that many mutation points exist in exon 1. The digestion products were then subjected to HRM analysis. The peaks of the PCR products from LpCKX1-F11R11 could not be detected by HRM analysis, as the denaturing

temperature of the DNA fragment in the reaction was over 99°C, and the highest temperature that Rotor-Gene Q can detect is 99°C (Appendix 4.10). Even though the machine could detect PCR products, the many mutation points in the screening region and the cleavage products could not be detected by HRM analysis.

In Trial 2, the template DNA was replaced by unmixed PCR products from individual genomic DNA samples. The cleavage products can be seen on the agarose gel (Figure 4.15B). The cleavage products in Trial 2 were subjected to HRM analysis, but as for Trial 1, the HRM peaks of the LpCKX1-F11R11 products could still not be detected.

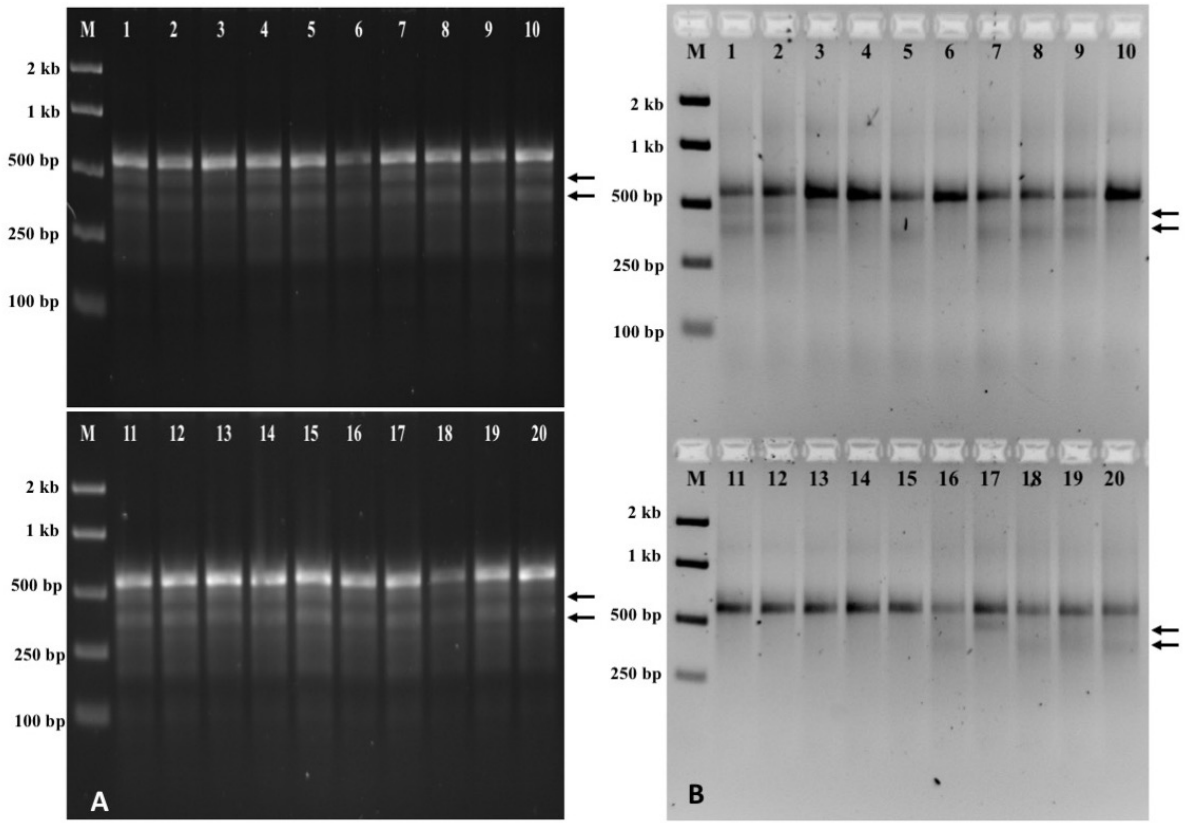


Figure 4.15 Cleavage products of exon 1 in *LpCKX1* from cv. Nui in Trial 1 (A) and Trial 2 (B). Details of Trial 1 and 2 refer to Table 4.8. The cleavage products are indicated by arrows. All PCR products were checked on a 3% (w/v) agarose gel with TAE buffer at 10 V/cm for 50 min. M: EasyLadder I.

4.4 Discussion

4.4.1 CEL I digestion with HRM analysis

TILLING or EcoTILLING has been successfully applied to a variety of crops to identify useful mutations, such as *TaMlo* and *waxy* in wheat (Slade et al., 2005; Acevedo-Garcia et al., 2017), and *OsBZIP* in rice (Till et al., 2007), as an efficient and high-throughput method. Identification of point mutations in the gene of interest is the core of the TILLING approach. Denaturing High Performance Liquid Chromatography (DHPLC) was used to detect point mutations in the original TILLING method (McCallum et al., 2000) and then an improved method with a lower-cost and greater efficiency was developed, combining enzyme digestion with CEL I nuclease to cleave mismatches in the double stranded DNA and gel electrophoresis with the LiCOR gel analyser (Colbert et al., 2001). To further improve the efficiency and reduce the cost of mutation detection, alternatives to the LiCOR gel based detection method have included non-denaturing polyacrylamide gel (Uauy et al., 2009), agarose gel (Dong et al., 2009b), or capillary electrophoresis (Perry et al., 2003; Stephenson et al., 2010). In recent years, HRM analysis has been applied to substitute gel electrophoresis of CEL I nuclease cleavage products in arabidopsis (Bush & Krysan, 2010), brassica (Lochlainn et al., 2011) and wheat (Dong et al., 2009b; Acevedo-Garcia et al., 2017). Even though HRM analysis is a sensitive method for detecting variations in sequence, such as SNPs, this technique requires PCR products with a short length of up to 350 bp for high resolution. In this chapter, HRM analysis was used as the alternative to the gel-based platform to check CEL I nuclease cleavage products.

CEL I nuclease was isolated from celery in the lab and the CEL I digestion conditions were optimised using different Taq DNA polymerases. CEL I was first discovered by Oleykowski et al. (1998), and they demonstrated that CEL I was able to detect a mismatched site in heteroduplexes. The purified CEL I showed more than one protein band in the range of 34-39 kDa (Oleykowski et al., 1998); However, the homemade CEL I showed two bands at 48 and 52 kDa, respectively (Figure 4.2). One possibility is that the homemade CEL I was not sufficiently purified, resulting in the CEL I nuclease binding with other proteins, and causing it to show a larger size on the gel.

Till et al. (2004) demonstrated that the pH and ion concentration in the PCR mixture, the digestion time and amount of CEL I in the digestion reaction can affect the cleavage results.

The three sets of master mixes showed smeared bands on the gel, but they also presented different efficiencies of CEL I digestion (Figure 4.4). With extended incubation time, the smeared bands became more noticeable (Figure 4.4). The smeared bands probably resulted from the degradation of DNA fragments, because CEL I nuclease has a 5'-3' exonuclease activity that cuts the ends of DNA (Qiu et al., 2004). Interestingly, regardless of how long the reaction was incubated, the Bioline Taq master mix showed similar smeared bands after 30 min incubation for, compared with Roche Taq master mix and the Type-it kit (Figure 4.4). It is possible that the Bioline Taq buffer was reducing the activity of exonuclease in CEL I as smearing was limited.

The EvaGreen dye was selected to be added into the CEL I digestion reaction to detect the cleavage products during the HRM analysis (Appendix 4.4). As shown, the EvaGreen dye may have had adverse effects on the CEL I digestion, as the cleavage products were hard to detect on the gel photo when EvaGreen dye was added to the reaction master before incubation (Figure 4.7). It is possible that EvaGreen dye could insert into the DNA groove and this insertion may affect the CEL I nuclease cleavage. However, the HRM analysis showed the same yield of cleavage products whether or not the EvaGreen dye was added (Figures 4.7 and 4.8). This result demonstrates that the EvaGreen dye did not affect the cleavage activity of CEL I, but it may have enhanced the 5'-3' exonuclease activity leading to an increase in degradation of DNA fragments. Meanwhile, compared to the HRM analysis of the PCR products, the denaturing temperature of the cleaved DNA fragments was significantly increased by CEL I and its digestion buffer (Figure 4.6).

Using the optimised conditions, the model DNA substrates from the hybridised PCR products of the G/C plasmids were cleaved by CEL I and the cleavage products were detected by HRM analysis (Figures 4.9 and 4.10). However, from the HRM analysis results, only one cleavage peak is shown on the melting curve, compared with the negative control without CEL I treatment or when the homozygotes were used as template. The two cleavage products should show as two small peaks on the melting curve. One possibility is that the two cleavage products from the heterozygote template shared similar denaturing temperatures, causing the peaks of the two cleavage products to overlap (Figure 4.10).

For the TILLING approach, CEL I nuclease has been used to screen mutations in PCR products from pooled individual genomic DNA (Colbert et al., 2001; Perry et al., 2003; Till et al., 2003). In arabidopsis, the pooling of genomic DNA from eight individuals, allowing

detection of 1 in 16 alleles, has been used to screen mutations (Till et al., 2003). Another study indicated that Surveyor, a CEL I nuclease commercial kit, could detect rare mutants present in as low as 1 in 32 copies (Qiu et al., 2004). The result presented in Figures 4.11 and 4.12 demonstrated that a mutation could be detected at 1 in 10 copies by gel electrophoresis and HRM analysis. For an optimal result on HRM analysis, the ratio of mixed alleles greater than 1 in 5 copies (the ratio of 2:8) is recommended, as the fluorescent signals of cleavage products were strongest when the ratio was greater than that (Figure 4.12).

4.4.2 Sequence determination of *LpCKX1* and *LpCKX2* genes

Both CKX1 and CKX2 have been implicated in determining seed size and/or seed number (Jameson & Song, 2016). The comparative genomics approach was used for identification of *CKX1* and *CKX2* in perennial ryegrass. Using the homologues of *CKX1* and *CKX2* from other crops (Ashikari et al., 2005a; Feng et al., 2008; Song et al., 2012), the full length cDNA from *LpCKX1* and the second half of *LpCKX2* were identified through BLAST searching the perennial ryegrass transcriptome database. Phylogenetic analysis showed that the putative *LpCKX1* and *LpCKX2* have a close relationship with their Poaceae homologues, suggesting that both genes might potentially have a similar function in perennial ryegrass (Figure 4.13).

Because seed size is more important than seed number for perennial ryegrass seed production (FAR, personal communication), the *LpCKX1* was selected as the target gene for mutation screening. From multiple alignment of the full length of *LpCKX1* (Appendix 4.7), the DNA sequences presented a high degree of genetic diversity. It was concluded from the study of genotyping with SSR markers (detailed in Appendix 4.1), that perennial ryegrass has a high degree of genetic diversity, which was not unexpected as it is an outcrossing species. However, the sequencing results of the full length *LpCKX1* gene showed that the coding regions also have a high degree of genetic diversity. This indicated that mutation screening in perennial ryegrass, irrespective of the approach, could be challenging.

The new method to detect CEL I digestion products with HRM analysis (Section 4.2.1), still needed to be tested in a practical mutation screening application. Exons 1 and 2 in *LpCKX1* were suitable for mutation screening because their length (about 650 bp), which is easy to detect using agarose gel electrophoresis after CEL I digestion. For the same reason, the exon 3 in *LpCKX1* (about 300 bp) and exons 4 and 5 in *LpSH1* (48 and 78 bp) were not long enough, not suitable for the CEL I+HRM method detection. From the multiple sequence

alignment of exon 2 of *LpCKX1* from 11 individuals of cv. Nui (Appendix 4.8), 19 SNPs were identified within the 578 bp of exon 2, about one SNP for every 31 base pairs. Consequently, exon 1 was selected as a screening region to test the new method, even though many SNPs also existed in exon 1, but there were fewer in exon 2.

4.4.3 Detection of mutations in *LpCKX1* using the optimised CEL I+HRM method

Two trials were designed to test CEL I+HRM as a method for mutation screening (Table 4.8), but they demonstrated that the new method was not suitable for mutation screening in perennial ryegrass. There are two main reasons. Firstly, many SNPs exist in exon 1 of *LpCKX1*, resulting in many cleavage bands visible on the agarose gel (Figure 4.15) and strong background noise detected by HRM analysis (Appendix 4.10). Secondly, the DNA sequences from *LpCKX1* have a high GC-content, and the presence of CEL I nuclease and CEL I digestion buffer in the reaction master increased the denaturing temperature of the DNA fragment, resulting in the Rotor-Gene Q machine not being able to detect the melting temperature (Appendix 4.10). Therefore, the results of both trials indicated that the CEL I+HRM method requires that the DNA sequence must have a denaturing temperature that can be detected by the machine, or that the CEL I digestion buffer needs to be further optimised. This method, which is clearly problematic with a genetically diverse species, could be used with species with high homozygosity, such as rice, wheat, and pea. Based on these two trials, the new method is not recommended for screening mutations in perennial ryegrass, because of the high degree of genetic diversity in perennial ryegrass and the high denaturing temperature of the PCR products from the CEL I digestion reaction master. Therefore, mutation screening in *LpSH1* and *LpCKX1* was carried out by the HRM analysis without CEL I enzyme.

4.4.4 Summary

A new method for detecting CEL I nuclease cleavage products using HRM analysis was optimised. The CEL I+HRM method was tested in a practical application of mutation screening, but the results suggest that this method is not suitable for a highly heterozygous species, such as perennial ryegrass, or a target gene with a high GC content, such as *LpCKX1*. Additionally, in contrast with exons 1 and 2 of *LpCKX1*, exon 3 appeared highly conserved as only two mutations were identified from the same 11 individuals for exon 2 sequencing, so this region was selected for mutation screening in Chapter 5.

5 Chapter 5: Mutation screening

5.1 Introduction

HRM analysis has been used to identify mutants from TILLING populations, such as *TaMlo* and *SBELLa* mutants in wheat (Dong et al., 2009b; Acevedo-Garcia et al., 2017) and *MEKK* mutant in arabidopsis (Bush & Krysan, 2010). Because HRM analysis requires the amplicon length to be less than 350 bp for accuracy (Reed & Wittwer, 2004), the selection of the region to screen for mutations is very important. For example, exon 9 of *TaMlo* was selected as the target for mutation screening in a wheat TILLING population, because the part of protein encoded by exon9 has been shown to confer powdery mildew susceptibility (Acevedo-Garcia et al., 2017). One mutation in this region, leading to an amino acid substitution, resulted in a loss-of-function of the protein, and enhanced resistance to powdery mildew disease. sixteen missense mutations of *TaMlo* were identified in the three wheat homoeologues, and each of the mutations had an altered amino acid leading to a functional change in the protein conferring resistance against powdery mildew in wheat (Acevedo-Garcia et al., 2017).

Based on the study in Chapter 4, CEL I+HRM method cannot be used in perennial ryegrass, but the HMR analysis on its own can be used for mutation screening in *LpSH1* and *LpCKX1*. In this chapter, seven DNA polymerases with EvaGreen fluorescent dye were tested to optimise a homemade HRM master mix.

Based on the experiments described in Chapter 3, the YABBY2 protein sequences from Poaceae and the LpSH1 protein sequence, shared high similarities with the zinc-finger domain and the YABBY domain (Figure 3.13), suggesting that these two domains could be functionally important. In a non-shattering sorghum cultivar, SC265, a splicing mutation occurs at the YABBY domain boarder, altering the YABBY domain coding region, and changing a shattering trait to a non-shattering (Lin et al., 2012). Additionally, mutations in *SH1* or other *YABBY* gene family members have been identified, such as one insertion (>4 kb) in *OsSh1* intron 3, one extremely large insertion (19.3 kb) in *ZmSh1-1* intron 1, an 83 bp insertion in *ZmSh1* exon 3 in two maize inbreds (Lin et al., 2012) and a 6-8 kb insertion in the *fas* intron 1 in domesticated tomato (Cong et al., 2008). All these mutations reduced either the transcript level or changed the YABBY domain structure and, consequently, changed the plant phenotype. This information suggests that two types of mutation might occur in perennial ryegrass: firstly, a long insertion in *LpSH1* which might reduce transcript level and

secondly, a small INDEL or SNP in *LpSH1* that might change the coding sequence. Therefore, the mutation screening in this work focused on exons 4 and 5 of *LpSH1*, as both exons cover the majority of the YABBY domain (refer to Figure 3.13).

As discussed in Chapter 1, both *CKX1* and *CKX2* have been implicated in determining in seed size and/or seed number in crops (Ashikari et al., 2005b; Song et al., 2012). According to Ashikari et al. (2005b), as 11-bp deletion was identified in exon 4 of *OsCKX2*, leading to a premature stop codon and a null mutation. This mutant produced over 400 seeds in the main panicle, significantly greater than the other two rice varieties in their study, about 164 and 306 seeds respectively. The full length sequence of *LpCKX1*and multiple alignment results described in Chapter 4 indicated that the DNA sequences spanning from intron 2 to exon 3 of *LpCKX1* were highly conserved in cv. Nui, compared with exons 1 and 2 (Appendices 4.7-4.9). In addition, exon 3 relates to the cytokinin binding domain in TaCKX2 in wheat (*Triticum aestivum*) (Feng et al., 2008). Therefore, this region from intron 2 to exon 3 was selected for mutation screening in this chapter.

5.2 Material and methods

5.2.1 Optimisation of homemade HRM master mixes

5.2.1.1 Preparation for HRM trial

HRM analysis is reported to be sufficiently sensitive to detect single nucleotide changes. Two pairs of plasmids differing in a single nucleotide were used to test commercial kits and homemade HRM master mix with different brands of DNA polymerases and EvaGreen dye.

One pair of plasmids (called G/C) was described in Section 4.2.1.2. Another pair of plasmids was constructed from two alleles of *LpSH1* which had one SNP difference. The alleles are described in Section 5.2.5.1 and named Allele4/8. The primer pairs for both plasmids were ordered from Macrogen Inc. and were optimised for specific PCR products. The details of the products from the primer pairs are listed in Table 5.1 and the primer sequences are listed in Table 5.2.

Table 5.1 Details of the PCR products of the two primer pairs. The SNP of each plasmid pair is highlighted in yellow and underlined.

Plasmid name	Amplicon sequences	Length (bp)	GC content	SNP position [a]
G/C	TCTATATAAGGAAGTTCATTTCAT TTGGAGAGGACACGCTCGAGATC ACAAGTTTGTACAAAAAAGCAGG CTTCATGAGACCAGAACGAAACCC CTTA(G/C)ATCTTAACAATTTGCCC GATGAGTACTCTAGAGATGGCAA ACAAGTCCTCGAAGA	154	41.1%	99nr G/C
Allele 4/8	GCAGCAGCTACCGAGAAGAGGCA GCGGGTTCCTTCAGCGTATAACCG ATTTATTAAGTAAGTTTCAATCAA TTAATTGC(G/A)AGCTTGTACTAAC AATAATTAAGCGGTATTGTCCTCT ACTAACC	124	41.1%	80nr G/A

a The SNP position is the sequence length from the 5´ end of the forward primer to the SNP site. Neither of the SNPs in the G/C plasmid or the Allele 4/8 is located in the primers. 99nr: there are 99 nucleotides from the start of the amplicon to the SNP site.

Table 5.2 Primer sequences for G/C plasmid and Allele4/8

Name	Sequences
GC-HRM-F1	TCTATATAAGGAAGTTCATTTCATTTGGAGAGGAC
GC-HRM-R1	TCTTCGAGGACTTGTTTGCCATCTCTAG
Allele4/8-HRM-F	GCAGCAGCTACCGAGAAGAG
Allele4/8-HRM-R	GGTTAGTAGAGGACAATACCGCTTA

5.2.1.2 Trial of HRM master mixes

All HRM analyses were performed using a Rotor-Gene Q (Qiagen, Hilden, Germany). The concentration of each template plasmid was diluted to about 50 pg/µl. A 5 µM primer mix consisted of 5 µM forward primer and 5 µM reverse primer. The homemade HRM master mixes were made from seven different DNA polymerases and EvaGreen Dye (20x in water, Cat. No. 31000, Biotium, USA). The details of each commercial HRM kit and DNA polymerase are listed in Table 5.3. The dNTPs mix in each reaction was made from dNTP set (Bioline, UK), and $MgCl_2$ was used from Bioline if the DNA polymerase reagent kit did not supply it. All reaction components and the cycling programs for HRM master mixes are attached in Appendix 5.1. Additionally, the primers, GC-HRM-F1/R1, with HPLC level of purification, were synthesised by Macrogen and used for HRM analysis with two commercial kits, Qiagen Type-it HRM kit and Biotium Fast EvaGreen qPCR Master Mix.

Table 5.3 Details of each commercial HRM kit and DNA polymerases.

	Name	Company	Country
Commercial kits	Type-it HRM PCR Kit	Qiagen	Germany
	Fast EvaGreen qPCR Master Mix	Biotium	USA
	Brilliant HRM Ultra-Fast Loci Master Mix	Agilent	USA
DNA polymerases used in homemade master mixes	AmpliTaq Gold 360	Applied Biosystems	USA
	Roche Taq	Roche	Germany
	Bioline Taq	Bioline	UK
	HotStarTaq Plus	Qiagen	Germany
	FastStart Taq	Roche	Germany
	GoTaq	Promega	USA
	DyNAzyme II	Thermo Scientific	USA

5.2.2 Plant materials and DNA extraction

Twenty individual plants from each of 28 ryegrass cultivars from different seed companies were grown for four to six weeks in pots outside at the University of Canterbury glasshouse complex during 2016-2017. Twenty individual seedlings of each cultivar were transplanted into small pots with slow release fertiliser commercial potting mix. Details of each cultivar are listed in Table 5.4. Leaf material from individual seedlings was collected and stored at -20ºC for genomic DNA extraction. An additional 47 leaf samples were collected from the field at Kimihia Research Centre, Lincoln, supplied by PGG Wrightson (not listed in Table 5.4), and the two cultivars of ryegrass seeds supplied from PGG Wrightson are their paternal cultivars, No. 6 and 7 in Table 5.4. In addition, No.7, Med line 1, is originally from the Mediterranean region and shows reduced seed shattering. In total, leaf material from 627 individual ryegrass plants was collected and their genomic DNA was extracted following the method described in Chapter 2.

Table 5.4 Details of each cultivar and genomic DNA sample number.

No.	Cultivar name	Endophyte	Company	Ploidy	DNA sample No.
1	Nui			D	1-20
2	RI009				21-40
3	Stellar	Nil	Seed Force	D	41-60
4	Elital	Nil	Seed Force	T	61-80
5	Glencar		DLF	T	81-100
6	Asset italian seed		PGG Wrightson		161-170
7	Med line 1		PGG Wrightson		181-200
8	Denver		DLF	D	201-220
9	Mezo		DLF	D	221-240
10	Alto	AR1	Agriseed	D	241-260
11	Rohan	NEA3	Agriseed	D	261-280
12	Arrow	AR1	Agriseed	D	281-300
13	Rohan	NEA2	Agriseed	D	301-320
14	Stellar	AR1	Seed Force	D	321-340
15	Meridian	AR1	Agriseed		341-360
16	Bronsyn	AR1	Agriseed	D	361-380
17	Trojan	NEA2	Agriseed	D	381-400
18	Bealey	NEA2	Agriseed	T	401-420
19	Basuto	AR1	DLF	T	421-440
20	Novello		DLF	T	441-460
21	Kentaur		DLF	T	461-480
22	MTX608				481-500
23	SAMSON	AR1	Agricom	D	501-520
24	Colosseum		PGG Wrightson		521-540
25	Arena		PGG Wrightson		541-560
26	R1313		FAR code		561-580
27	Halo		Agricom	T	581-600
28	CERES ONE50		Agricom	D	601-650

5.2.3 First run of nested PCR for mutation screening of *LpSH1* and *LpCKX1*

For HRM analysis, the nested PCR approach was applied to screen mutations in *LpSH1* and *LpCKX1*. The first run of nested PCR was a standard PCR amplification and the primers were designed and optimised (Table 5.5). The selected primers, LpSH1-seq-F1R1 and LpCKX1-F5R14, were used for the first run of nested PCR with genomic DNA from individual ryegrass plants. The PCRs were performed with Platinum Taq DNA polymerase (Invitrogen, USA) and the reaction master mixes for both genes are listed in Table 5.6. The reaction was subjected to 95°C for 2 min, followed by 40 cycles of 95°C for 30 s, 58°C for *LpSH1* or 60°C for *LpCKX1*, annealing for 90 s, and 72°C extension for 40 s, and final 72°C extension for 10 min. The PCR products were then diluted about 300-fold with Milli-Q water as PCR templates for the second run of nested PCR and HRM analysis.

Table 5.5 Primer sequences of *LpSH1* and *LpCKX1* for the first round of nested PCR

Gene	Primer name	Sequences [a]
LpSH1	LpSH1-seq-F1	GCATGACGGATTAAGAGTATGCAGTTT
	LpSH1-seq-F2	TGACTGTTGACCTTGCCATATGTG
	LpSH1-seq-R1	ACTATGGGTTATGGTCAAGTGTGGAAA
	LpSH1-seq-R2	ATGGGTTATGGTCAAGTGTGGAAACT
	LpSH1-seq-R3	TTTGACGGACGAAACTGGAGTAGT
LpCKX1	LpCKX1-F5	CCCGTGGCTCAACTTGCTCGTG
	LpCKX1-F6	TCCCGCATCGCCGACTTCGAC
	LpCKX1-F12	CTTCGGCGGCATCCTCCAG
	LpCKX1-F13	CGTCTTCGGCGGCATCCTC
	LpCKX1-R14	AGTCTAGTTGAAGATGTCCTGTCCTG
	LpCKX1-R15	TCAGTCTAGTTGAAGATGTCCTGTCC
	LpCKX1-R1	CAGTCTAGTTGAAGATGTCCTGTCCT
	LpCKX1-R2	GTCCCACTTGCCGCCGAAAT

[a] Expected product size and annealing temperature for the selected primers are shown in Table 5.8.

Table 5.6 Components of the first run PCR master mixes for *LpSH1* and *LpCKX1*.

	LpSH1	*LpCKX1*
Components	Volume per 10 µl reaction (µl)	
Platinum Taq DNA polymerase	0.04	0.04
Platinum 10x PCR buffer	1	1
50 mM $MgCl_2$	0.3	0.3
20 mM dNTP	0.1	0.1
KB Extender	0	0.5
10 mM Primer	0.7	0.7
Template DNA	1	1
PCR-grade water	Make up to 10	

5.2.4 Second run PCR and HRM analysis

The second run of nested PCR was to screen the different alleles with HRM analysis. Two small regions within the first amplicon of each gene were screened separately (Figure 5.1). Specific primers for HRM analysis were designed and optimised (Table 5.7). The selected primers including LpSH1-HRM-F2R2 (*LpSH1* amplicon 1), LpSH1-HRM-F7R5 (*LpSH1* amplicon 2), LpCKX1-HRM-F13R7 (*LpCKX1* amplicon 1), LpCKX-HRM-F4R5 (*LpCKX1* amplicon 2) (Figure 5.1), were used for HRM analysis with the templates from diluted PCR products of the first nested PCR run. All HRM analyses were performed with EvaGreen dye (20x in water, Biotium, USA) using a Rotor-Gene Q (Qiagen, Hilden, Germany). Each HRM reaction contained 0.375 Units of AmpliTaq Gold 360 DNA Polymerase (Applied Biosystems, USA), 1.5 µl of 10x AmpliTaq PCR buffer, 0.2 mM of each dNTP (Bioline,

UK), 1.8 µl of 25 mM $MgCl_2$, 0.75 µl of 20x EvaGreen dye, 0.33 µM of each primer and 1 µl of template DNA in a total volume of 15 µl. The reaction was subjected to 94°C for 10 min, followed by 40 cycles of 95°C denaturing for 30 s, 55°C annealing and extension for 1 min. The HRM analysis was performed as follows: 60°C for 5 min, and then the temperature was increased from 70 to 95°C at a rate of 0.1°C/2sec.

With each HRM analysis, several samples from cv. Nui were added as controls. These samples were sequenced and shown to have different alleles. If any other samples presented the same patterns as the controls in the HRM analysis, this indicated that they had the same PCR products.

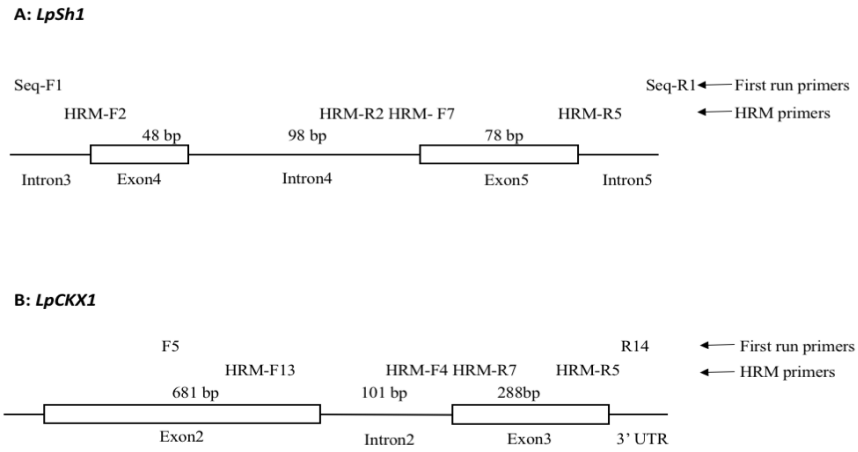


Figure 5.1 Structures of the gene region to screen mutations in (A) *LpSH1* and (B) *LpCKX1*, and primers location at approximate primers sites.

Table 5.7 Primer sequences for mutation screening in *LpSH1* and *LpCKX1* for HRM analysis.

	Name	Sequences [a]
LpSH1 amplicon 1	LpSH1-HRM-F1	ATGAGGATGAATCTTTGATACTGCA
	LpSH1-HRM-F2	GCAGCAGCTACCGAGAAGAG
	LpSH1-HRM-F3	TAAGCGGTATTGTCCTCTACTAACC
	LpSH1-HRM-R1	ATAGGTGTTATGCGTAGCATGTACT
	LpSH1-HRM-R2	GGTTAGTAGAGGACAATACCGCTTA
	LpSH1-HRM-F4	GCAGCAGCTACCGAGAAGAG
	LpSH1-HRM-F5	ATGAGGATGAATCTTTGATACTGCA
	LpSH1-HRM-R3	AGCTTGCATTTAATTGATTGAAACTTAC
	LpSH1-HRM-R4	GCTTGCATTTAATTGATTGAAACTTACTT
LpSH1 amplicon 2	LpSH1-HRM-F6	CCTCTACTAACCATACGATTCTGACACATA
	LpSH1-HRM-F7	CTAACCATACGATTCTGACACATAAATTTC
	LpSH1-HRM-R5	CATAGGTGTTATGCGTAGCATGTACTAAC
	LpSH1-HRM-R6	ATAGGTGTTATGCGTAGCATGTACT
LpCKX1 amplicon 1	LpCKX1-HRM-F1	ATCCTCCAGGGCACCGACAT
	LpCKX1-HRM-F2	TCGCTGGGCCTATGGTCATCT
	LpCKX1-HRM-F12	CTTCGGCGGCATCCTCCAG
	LpCKX1-HRM-F13	CGTCTTCGGCGGCATCCTC
	LpCKX1-HRM-R1	AGCATCGACACCGCGTAGAAC
	LpCKX1-HRM-R2	ACACCGCGTAGAACACCTCCT
	LpCKX1-HRM-R6	GCCACCGACGAGAAGAGCAT
	LpCKX1-HRM-R7	CGCCACCGACGAGAAGAGC
LpCKX1 amplicon 2	LpCKX1-HRM-F3	CGGTGGCGAACGACCTGAA
	LpCKX1-HRM-R3	AGTCTAGTTGAAGATGTCCTGTCCTG
	LpCKX1-HRM-R4	TTGAAGATGTCCTGTCCTGGAGAG
	LpCKX1-HRM-F4	CGGCGGAGGAGGTGTTCTAC
	LpCKX1-HRM-R5	GGTCGTACTTGTCCTTCATCTGG

[a] Expected product size and annealing temperature for primer selected are shown in Table 5.9.

5.2.5 Sequencing and sequence analysis

5.2.5.1 *LpSH1*

Twenty-one individuals with different melting curve patterns in the HRM analysis were selected (Table 5.8). PCR was performed with the genomic DNA from the 21 selected individuals, under the same conditions as the first nested PCR run (Section 5.2.2). The PCR products from primer LpSH1-seq-F1R1 were cloned using the TOPO TA Cloning Kit (Invitrogen, USA) or pGEM-T Easy Vector Systems (Promega, USA) following the user's manuals. The constructed plasmids were transformed into *E. Coli* DH5α competent cells. The plasmids were isolated from transformed cells using DNA-spin Plasmid DNA Purification kit (iNtRON Biotechnology Inc. South Korea) following the manual. All plasmids were sent to Macrogen for sequencing. The sequencing results were analysed by using multiple alignment to detect the mutation point in each allele sequence.

Table 5.8 Individuals selected for sequencing in *LpSH1* and *LpCKX1* based on genotype.

	Plant individual	Primer	Amplicon size (bp)
LpSH1	2, 4, 10, 12, 15, 21, 31, 36, 90, 146, 169, 182, 186, 193, 397, 403, 433, 489, 504, 566, 599	LpSH1-seq-F1R1	560
LpCKX1	23, 27, 301, 307, 361, 373, 509, 601	LpCKX1-F5R14	525

5.2.5.2 *LpCKX1*

Eight individuals with a single melting curve peak on the HRM analysis were selected for sequencing (Table 5.8). The genomic DNA from eight individuals was subjected to the first run of nested PCR using Platinum Taq DNA polymerase (Invitrogen, USA) and the PCR products were then separated on agarose gel with TAE buffer, at 7 V/cm for 40 min. The bands were collected and purified using UltraClean 15 DNA Purification Kit. The gel purified PCR products were then sequenced by Macrogen. The sequencing results were analysed by using multiple alignment to detect the mutation point in each allele sequence.

5.2.5.3 Sequence analysis of *LpCKX1* and *LpSH1*

For multiple sequence alignment, the exon and intron regions were aligned. The coding sequences of *LpSH1* and *LpCKX1* were assembled and realigned. After alignment of coding sequences, the translated amino acid sequences of both genes were analysed by using T-Coffee multiple sequence alignment server (http://tcoffee.crg.cat/).

5.3 Results

5.3.1 Trial of homemade HRM master mix

Two pairs of plasmids, G/C and allele4/8, were used to test a variety of homemade HRM master mixes, in comparison with commercial HRM kits. For the G/C plasmids, all homemade HRM master mixes and commercial HRM kits differentiated between homozygous and heterozygous sequences (G/G, C/C vs. G/C), whereas the homozygous sequences (G/G vs. C/C) cannot be discriminated. The HRM normalised fluorescence graphs from the Biotium HRM kit and AmpliTaq homemade HRM master mix are shown in Figure 5.2. The primers GC-HRM-F1/R1with HPLC level of purification were tested with the Type-

it HRM kit and the Biotium HRM kit, but the results were the same as with standard purification primers.

For allele4/8, a pair of plasmids with one SNP difference (G/A), five homemade master mixes, including AmpliTaq, DyNAzyme, Roche Taq, Roche FastStar Taq, and HotStar Taq and two commercial HRM kits, Biotium and Agilent, could discriminate between all three templates (4/4, 8/8, 4/8). In contrast, two homemade master mixes, with Bioline Taq and GoTaq and the Type-it HRM kit could not distinguish between the 4/4 and 4/8 templates. The HRM normalised fluorescence graphs from the Agilent HRM kit and AmpliTaq homemade HRM master mix are shown in Figure 5.3. Taken together, the AmpliTaq could be used for homemade HRM master mix.

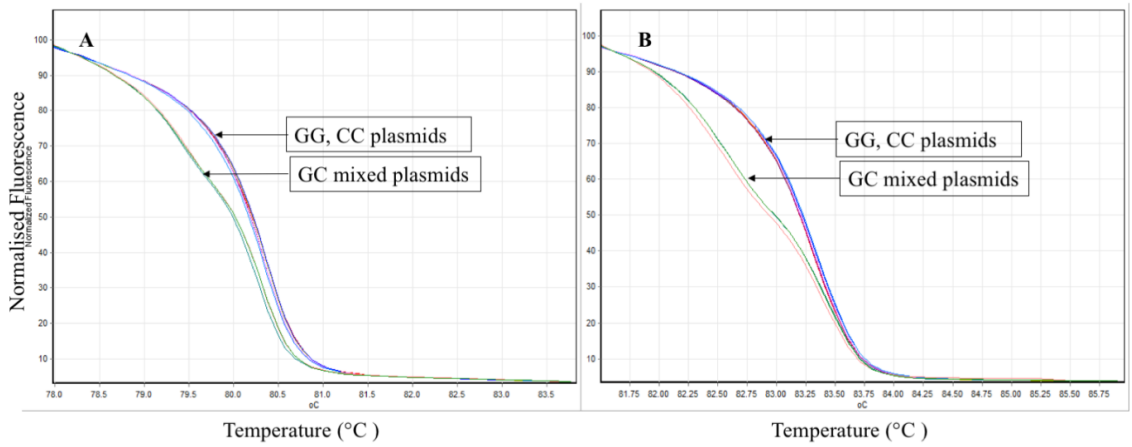


Figure 5.2 HRM normalised fluorescence graphs of GC-HRM-F1/R1 performed by the Biotium HRM Kit (A) and AmpliTaq homemade master mix (B).

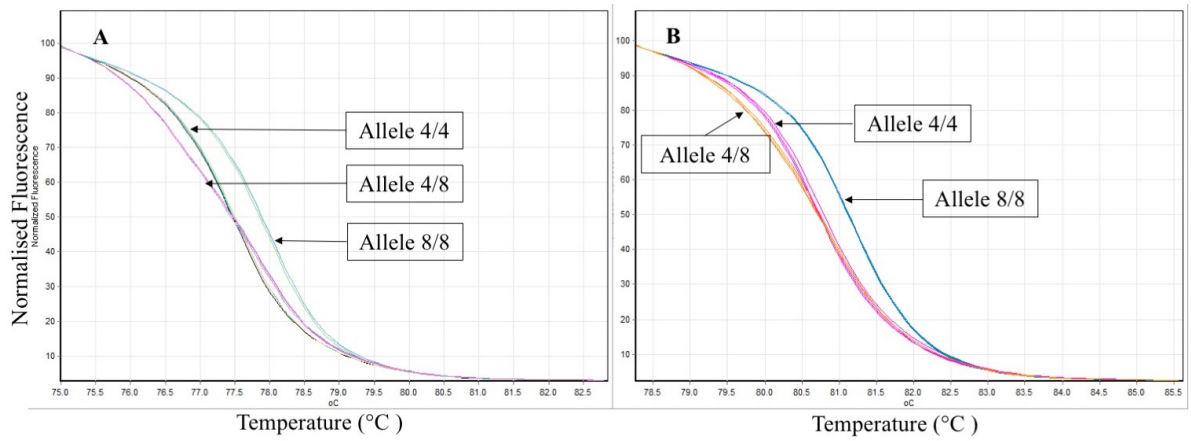


Figure 5.3 HRM normalised fluorescence graphs of Allele4/8-HRM-F/R performed by the Agilent HRM Kit (A) and AmpliTaq homemade master mix (B).

5.3.2 Mutation screening in *LpSH1*

The nested PCR approach was used to increase the resolution of the HRM analysis, in order to reduce non-specific binding in products due to the amplification of unexpected primer binding sites. For the *LpSH1* gene, the first run PCR was to amplify the region from LpSH1-Seq-F1 to LpSH1-Seq-R1, which contained the target region for mutation screening (Figure 5.1). For HRM analysis, two pairs of primers were selected from Table 5.7, as their PCR products have a sharp peak on the melting curve. Both amplicons spanned either the whole exon 4 and the first half of the intron 4 (LpSH1-HRM-F2R2), or the second half of intron 4 and the whole exon 5 (LpSH1-HRM-F7R5), as shown in Figure 5.1.

After HRM analysis, all samples with the same patterns were classified into one group, assuming that one pattern represented one individual. Each individual could contain 1 or 2 alleles if it was a diploid cultivar and 1 to 4 alleles if it was a tetraploid cultivar. For the following sequencing step, one individual was selected from each pattern. From the mutation screening results, 15 patterns were detected using the primer LpSH1-HRM-F2R2 and eight patterns were detected using primer LpSH1-HRM-F7R5 (Table 5.9; the pattern details of all individuals are in Appendix 5.2). In total, 21 individuals were selected for sequencing, each containing a different pattern (Table 5.8).

Table 5.9 Pattern number for each primer pair following HRM analysis

HRM Primer pairs	Amplicon size (bp)	Pattern No.	Genotype No.
LpSH1-HRM-F2R2	124	15	21
LpSH1-HRM -F7R5	137	8	
LpCKX1-HRM-F13R7	251	>40	N/A [a]
LpCKX1-HRM-F4R5	219	>40	

[a] The number of genotypes cannot be counted, as many melting curve patterns were identified from each HRM primer pair.

5.3.3 Mutation screening in *LpCKX1*

The nested PCR approach was also used for *LpCKX1*. For mutation screening, two primer pairs, LpCKX1-HRM-F13R7 and LpCKX1-HRM-F4R5, were selected for HRM analysis from Table 5.7. The primers for nested PCR and HRM analysis are indicated in Figure 5.1.

The mutation screening of *LpCKX1* with different cultivars gave a variety of melting curve patterns. The normalised fluorescent graphs and melt curves for primer LpCKX1-HRM-F4R5, and three ryegrass cultivars including diploid cv. Nui, cv. RI009 and cv. Stellar, are shown in Figure 5.4. Twenty individuals of cv. Nui had only two PCR products, while more than ten patterns were identified from 20 individuals for cv. RI009. After analysing the HRM results, at least 40 patterns were identified from all the ryegrass individuals and both *LpCKX1* primer pairs (Table 5.9).

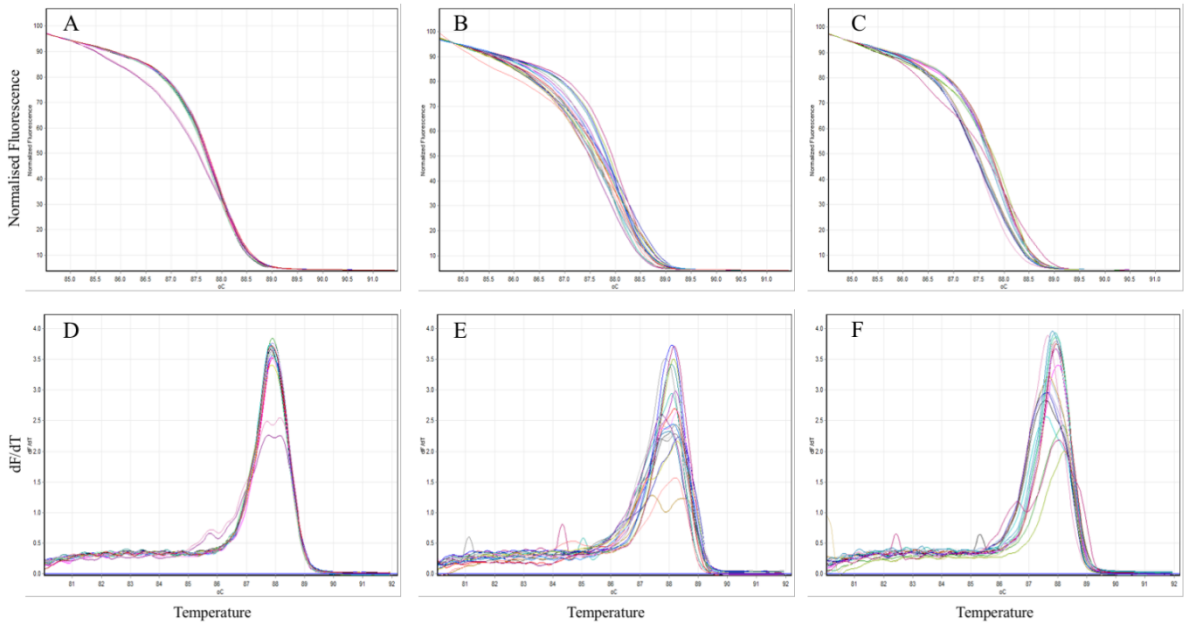


Figure 5.4 The normalised fluorescence (A, B, C) and melting curves (D, E, F) of LpCKX1-HRM-F4R5 from three diploid ryegrass cultivars, including cv. Nui (A and D), cv. RI009 (B and E); and cv. Stellar (C and F). Each line in the graph represents one amplicon from an individual seedling.

5.3.4 Sequencing of *LpSH1* and *LpCKX1*

For *LpSH1*, 21 genotypes were determined from the EcoTILLING population using two HRM primer pairs, and 19 of them were sequenced successfully in the region from primer LpSH1-seq-F1 to LpSH1-seq-R1 (Figure 5.1). After sequencing and alignment, 13 alleles with different DNA sequences were aligned and are shown in Figure 5.5. Based on the multiple alignment, many SNPs and INDELs exist in the introns, especially in intron 4. Focusing on exons 4 and 5 of each *LpSH1* allele, the number of different allele sequences reduced to nine, and seven SNPs were identified (Figure 5.6). The nine alleles with different DNA sequences were translated to four different amino acid sequences. The alignment result is shown in Figure 5.7.

For *LpCKX1*, many patterns were identified from the normalised fluorescence graphs (Figure 5.4), which means that many genotypes existed in the region identified for mutation screening and these different genotypes contained different numbers of allele sequences, so it was hard to sequence all of them using the TA-cloning technique. Therefore, eight samples with a single peak on the melt curve were selected for sequencing, as a single peak

represented one PCR product. The PCR products from primer LpCKX1-F5R14 were sequenced. Only six individuals (DNA sample no. 27, 301, 307, 361, 509, 601, refer to Tables 5.4 and 5.8), were sequenced successfully. From the alignment result (Figure 5.8), many SNPs and INDELs exist in the six alleles, but most mutations were present in the intron region, as noted also with *LpSH1*. To analyse the coding region of *LpCKX1*, intron 2 in each allele was removed and the exon parts were assembled and realigned (Figure 5.9). Two sequenced PCR products, amplified using primer LpCKX1-F13R15 and genomic DNA from Nui 3 and 14 (refer to Section 4.2.2.1), were also added for coding region realignment. The PCR products from Nui 3 and 14 were amplified by different primer pairs, so the flanking sequences of the two PCR amplicons were not aligned, but the majority of the DNA sequences overlapped. The realignment of the coding regions of the eight DNA sequences (Figure 5.9), showed that there are only six different alleles, as the alleles from 301, 307, 361 share the same DNA sequence within the exon. The six different allele sequences have eight SNPs. The alleles with different DNA sequences were translated to amino acids and aligned (Figure 5.10). Alleles from 301, 307, 361, 601 translated to the same amino acid sequence. From the alignment of amino acid sequences, five different amino acid sequences were identified.

The amino acid sequences translated from parts of *LpSH1* and *LpCKX1* were analysed by using the T-COFFEE multiple sequence alignment server. None of the translated amino acid sequences presented a functional change. The results were downloaded and are in Appendices 5.3 and 5.4.

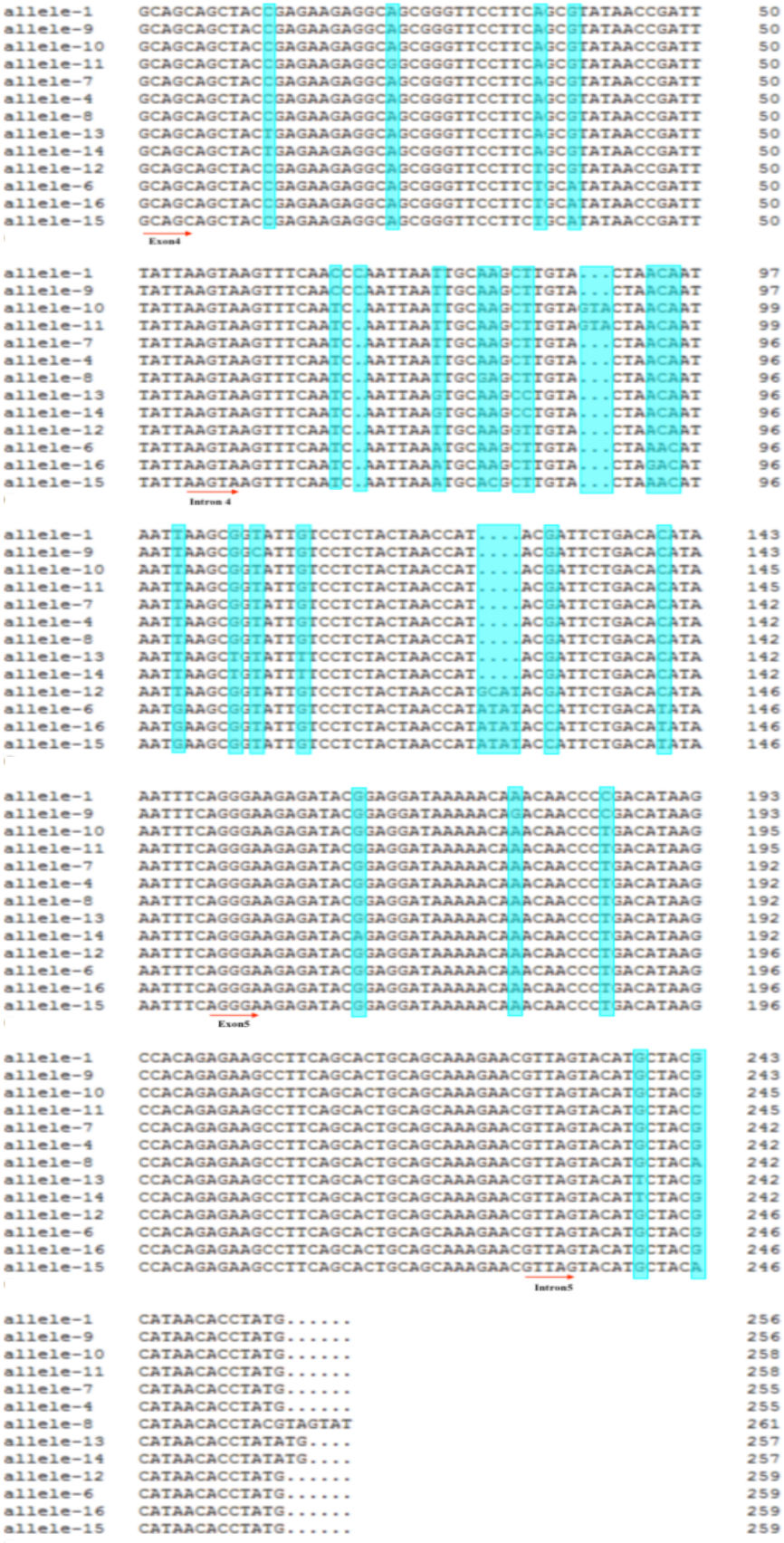


Figure 5.5 Alignment of sequences from mutation screening region in *LpSH1*. The SNPs and INDEL are highlighted in blue, and the exons and introns are indicated by red arrows.

Figure 5.6 The alignment of the coding region of *LpSH1*. Seven SNPs are highlighted in blue, and the exons 4 and 5 are underlined by red and blue.

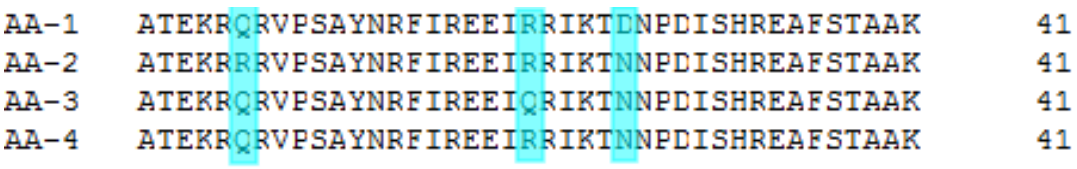


Figure 5.7 The alignment of translated amino acid sequences from exons 4 and 5 in *LpSH1*. The four sequences were translated from the alleles aligned in Figure 5.6. AA-1: allele-9; AA-2: allele-11; AA-3: allele-14; AA-4: alleles 1, 4, 6, 7, 8, 10, 12, 13, 15, and 16. Three sites of amino acid change are highlighted by blue.

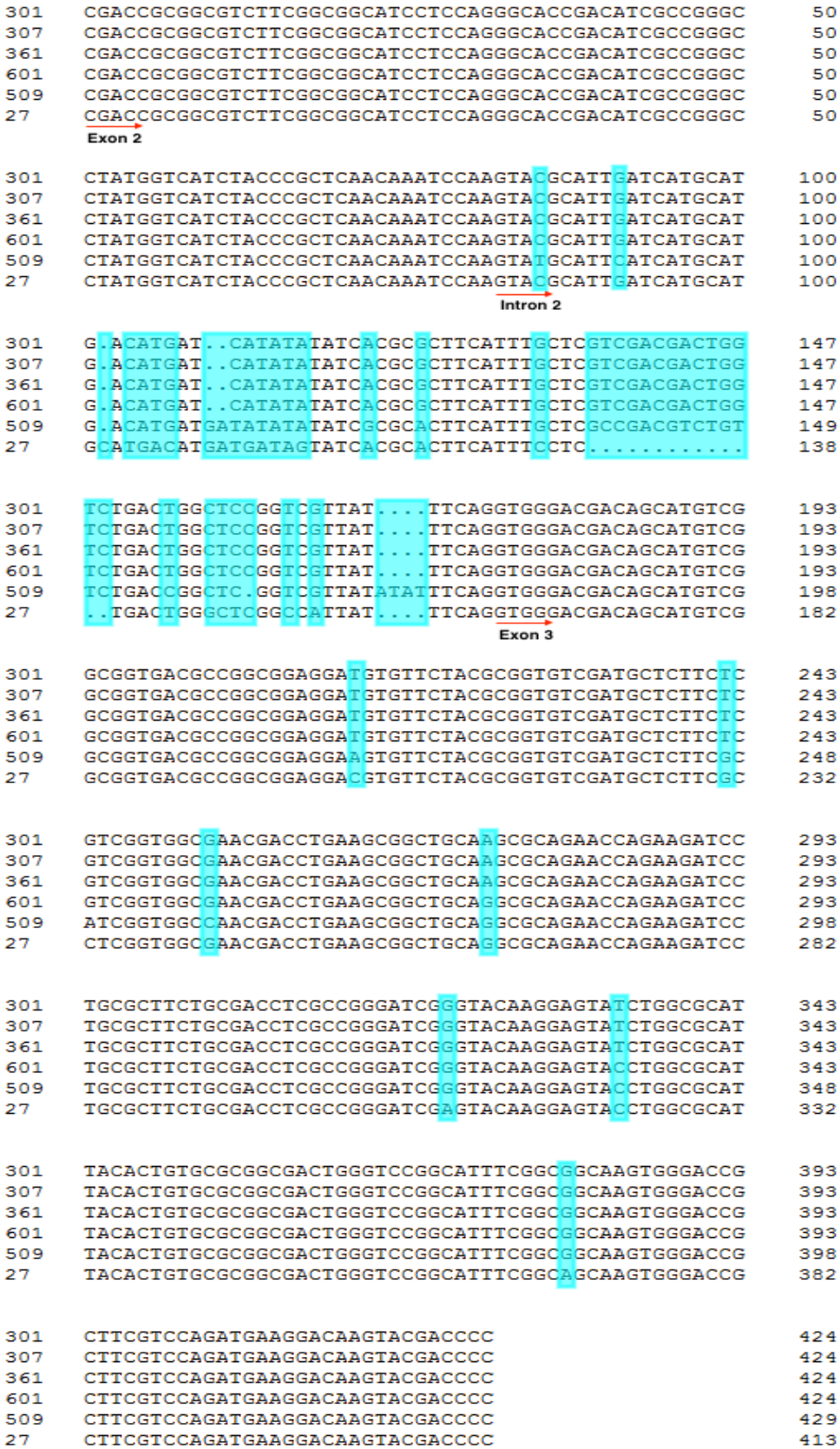


Figure 5.8 Alignment of sequences from mutation screening region in *LpCKX1*. The SNPs and INDELs are highlighted in blue, and the exons 2 and 3 and intron 2 are indicated by red arrows.

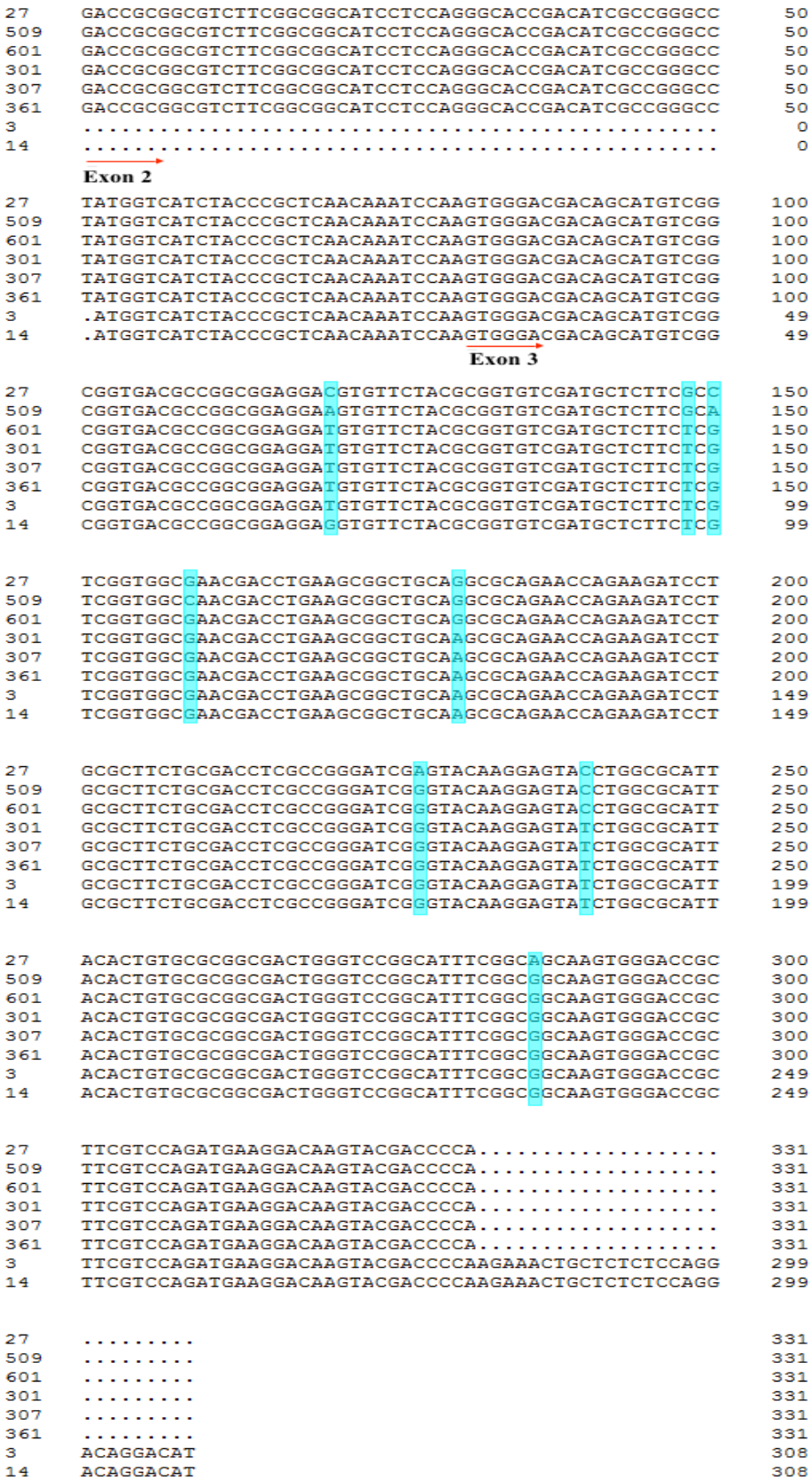


Figure 5.9 Multiple alignment of part of the coding region of *LpCKX1*. The alleles from cultivars No. 3, 301,307, 361 have the same DNA sequences in the coding region. The eight SNPs are highlighted in blue, and exons 2 and 3 are indicated by red arrows.

```
AA-1    DRGVFGGILQGTDIAGPMVIYPLNKSKWDDSMSAVTPAEDVFYAVSMLFA     50
AA-3    DRGVFGGILQGTDIAGPMVIYPLNKSKWDDSMSAVTPAEEVFYAVSMLFA     50
AA-2    DRGVFGGILQGTDIAGPMVIYPLNKSKWDDSMSAVTPAEDVFYAVSMLFS     50
AA-4    ..................MVIYPLNKSKWDDSMSAVTPAEDVFYAVSMLFS     33
AA-5    ..................MVIYPLNKSKWDDSMSAVTPAEEVFYAVSMLFS     33

AA-1    SVANDLKRLQAQNQKILRFCDLAGIEYKEYLAHYTVRGDWVRHFGSKWDR    100
AA-3    SVANDLKRLQAQNQKILRFCDLAGIGYKEYLAHYTVRGDWVRHFGGKWDR    100
AA-2    SVANDLKRLQAQNQKILRFCDLAGIGYKEYLAHYTVRGDWVRHFGGKWDR    100
AA-4    SVANDLKRLQAQNQKILRFCDLAGIGYKEYLAHYTVRGDWVRHFGGKWDR     83
AA-5    SVANDLKRLQAQNQKILRFCDLAGIGYKEYLAHYTVRGDWVRHFGGKWDR     83

AA-1    FVQMKDKYDP.........                                   110
AA-3    FVQMKDKYDP.........                                   110
AA-2    FVQMKDKYDP.........                                   110
AA-4    FVQMKDKYDPKKLLSPGQD                                   102
AA-5    FVQMKDKYDPKKLLSPGQD                                   102
```

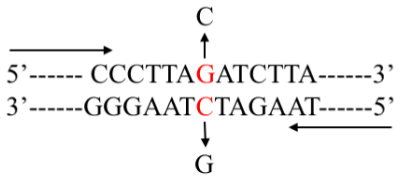
Figure 5.10 Multiple alignment of translated amino acid sequences from exons 2 and 3 in *LpCKX1*. The five sequences were translated from the alleles aligned in Figure 5.9. AA-1: allele-27; AA-2: allele-301, 307, 361 and 601; AA-3: allele-509; AA-4: allele-3; AA-5: allele-14. The four sites of amino acid change are highlighted by blue.

5.4 Discussion

5.4.1 Trial of a variety of HRM master mix

The aim of this project was to identify useful alleles from a perennial ryegrass EcoTILLING population. HRM analysis was used to screen mutations in the target genes of interest. This method is relatively sensitive for detecting SNPs, but it requires high specificity PCR amplification, and the PCR amplicons to have an ideal length (<350 bp). The selection of quality DNA polymerase is essential for HRM analysis. A variety of DNA polymerases combined with EvaGreen dye and commercial HRM kits were compared using two DNA substrates. From the HRM analysis of each master mix, the homozygotes of G/C plasmids could not be differentiated using any of the master mixes (e.g. Figure 5.2). Because the homozygous SNP genotypes differed only by melting temperature, if instruments cannot detect the difference between the two alleles, then these two genotypes would not be able to be distinguished. Four classes of SNP were grouped by their combinations of bases (Class 1: C/T & G/A; Class 2 C/A & G/T; Class 3: C/G; Class4: A/T). The melting temperature of the homozygotes in Classes 3 and 4 is less than 0.5°C (Liew et al., 2004). Also, the difference between amplicon melting temperatures is also affected by the nearest-neighbour bases next to the SNP site (Liew et al., 2004). In the G/C plasmids, the different homozygotes could not be differentiated by dissociation temperature using HRM analysis, because the SNP belongs to Class 3 and the nearby bases are the same (Figure 5.11A). This situation leads to the two homozygotes only having a simple inversion and nearest-neighbour stability, so the dissociation temperatures for the two homozygotes are almost identical. In contrast, the two homozygotes were distinguished easily in the other plasmid pair, allele4/8 (Figure 5.11 B), because of the SNP class and the bases neighbouring the SNP site.

A. GC plasmids (Class 3)



B. Allele 4/8 (Class 1)

Figure 5.11 Details of the SNPs in (A) the G/C plasmid and (B) the allele 4/8.

Even though the homozygotes of the G/C plasmids could not be distinguished, most of the homemade master mixes and commercial HRM kits successfully differentiated the homozygotes and heterozygotes in alleles4/8. Finally, the AmpliTaq in a homemade master mix is suggested as optimal for PCR amplification followed by HRM analysis with EvaGreen dye, because the PCR products amplified by the homemade master mix with AmpliTaq show sharp peaks in melting curves, compared with other master mixes.

5.4.2 Mutation screening

Genomic DNA was isolated from 20 individuals from each of 30 cultivars most of which are widely used New Zealand cultivars. Pembleton et al. (2016) constructed a Neighbour-Joining tree mainly containing perennial ryegrass and Italian ryegrass, which is genotyped by using 296 SNP marker loci. Nine cultivars used in this project were represented in this phylogenetic tree and located at different clusters, suggesting that the cultivars or the EcoTILLING population used in this thesis contained a wide genetic background.

The results of the HRM analysis showed that the different patterns in the normalised fluorescence graphs could be identified from different individuals (e.g. Figure 5.4). Each pattern represented one genotype, and each genotype probably contained different numbers of alleles: one or two alleles in a diploid cultivar and one to four alleles in a tetraploid cultivar. After sequencing of the amplicons with different patterns, the different alleles of the

target genes were analysed. The amino acid sequences from each different allele were translated using MEGA 7 (Figures 5.7 and 5.10). The amino acids translated from the identified alleles of both genes were analysed. Unfortunately, no amino acid changes were detected that could potentially alter the protein structure and gene function.

For the *LpSH1* gene, 21 genotypes were identified from the EcoTILLING population using two HRM primers from the mutation screening region. From the results of pattern analysis, the diploid cultivars, either commercial cultivar or ecotypes, shared similar patterns, but several patterns in the tetraploid cultivars were not identified in any of the diploid cultivars. One explanation for this is that the tetraploid cultivars contain more alleles, so there is more chance to combine different alleles, and therefore presenting different patterns in melting curve analysis. The cultivar Med line 1, supplied by PGG Wrightson, originated from the Mediterranean region. It showed several genotypes, which did not exist in any other cultivars (Appendix 5.2). This result suggests that cultivar Med line 1 is genetically distinct from the cultivars collected from New Zealand.

For *LpCKX1*, the number of patterns in the normalised fluorescence or melting curves using LpCKX1-hrm-F4R5 varied among different cultivars (Figure 5.2). The PCR products of LpCKX1-hrm-F4R5 were amplified from exon 3 of *LpCKX1*. Based on the full length multiple alignment, the DNA sequence in exon 3 of *LpCKX1* showed a high conservation (refer to Section 4.2.2.1 and Appendix 4.9). In contrast, the HRM analysis with primer LpCKX1-hrm-F4R5 revealed a variety of melting curves patterns among different cultivars, meaning that exon 3 was not a conserved region. One explanation for this is that different perennial ryegrass cultivars have various degrees of genetic diversity, because some commercial cultivars were bred from different numbers of inbred lines, resulting in some cultivars having high levels of heterozygosity (Momotaz et al., 2004). This result is also supported by the genotyping using SSR markers of the ryegrass cultivars, which showed different degree of genetic diversity in different cultivars (Appendix 4.1). This indicates that exon 3 of *LpCKX1* gene also has a high degree of genetic diversity, similar to exons 1 and 2 in *LpCKX1*.

Comparing the mutation screening results of the two target genes, the number of patterns from each primer pair was also very varied as shown in Table 5.9. There are 15 and 8 patterns identified from two primer pairs of *LpSH1* (Table 5.9 and Appendix 5.2), significantly fewer than the over 40 patterns from both primer pairs for *LpCKX1* (Table 5.9).

Even though the length of amplicons from *LpCKX1* (about 200 bp), are longer than that from *LpSH1* (about 120 bp), this does not explain why *LpCKX1* has many more alleles than *LpSH1*.

A flow chart of the mutation screening process is shown in Figure 5.12. From the sequencing results, six alleles of *LpCKX1* were identified from eight individuals, whereas only 13 alleles of *LpSH1* were identified from all 627 individuals of the 31 ryegrass cultivars. Finally, five amino acid sequences were identified in *LpCKX1* (Figure 5.10) and four amino acid sequences were identified in *LpSH1* (Figure 5.7). Taken together, the two target genes have a different degree of genetic diversity, with the implication that *LpSH1* is more conserved than *LpCKX1*. This suggests that seed shattering is under stronger selection than seed size or number. This agrees with the work of Tang et al. (2013), which indicated that the seed shattering phenotype is often under opposing selection in crops and their wild relative, because non-shattering may be selected by domestication and shattering may be selected by wild species for seed dispersal.

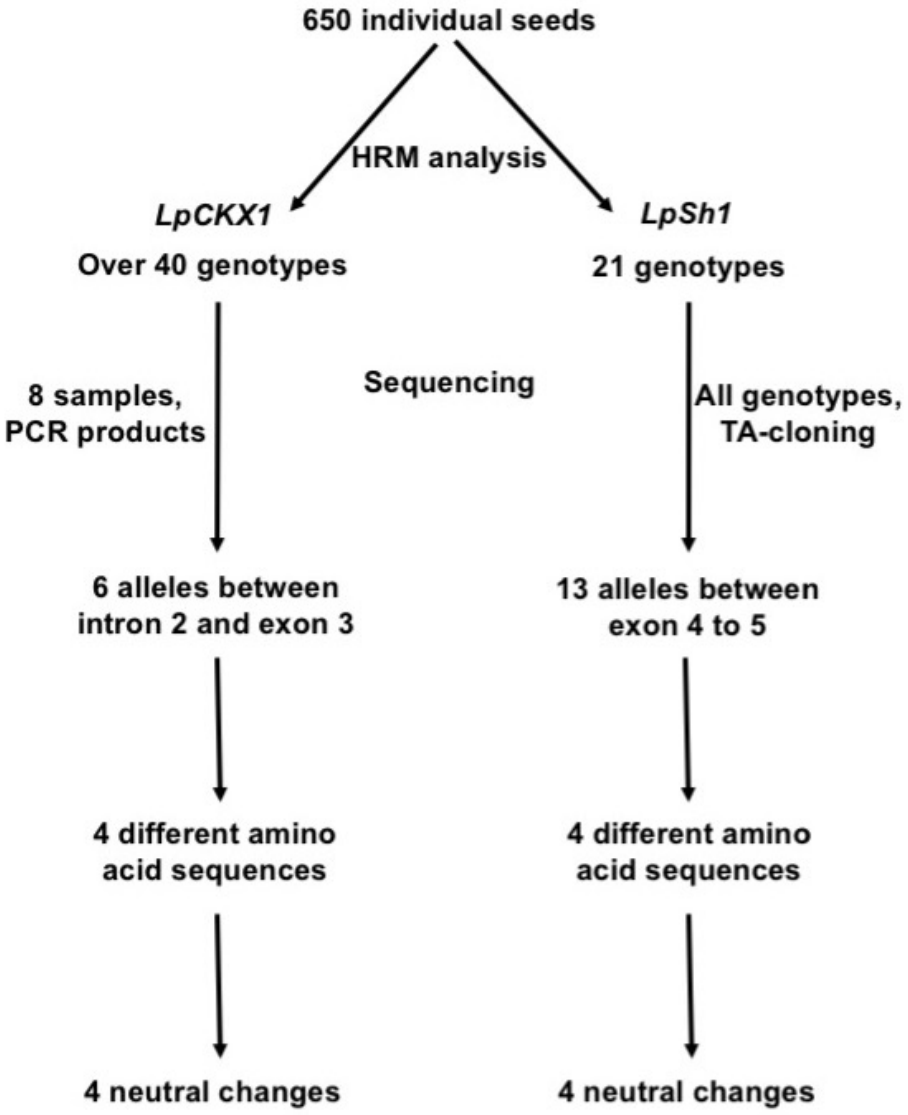


Figure 5.12 The flow chart of the mutation screening process used in this project. Genomic DNA was extracted from 650 individuals from 30 cultivars. After mutation screening using HRM analysis, over 40 genotypes of *LpCKX1* and 21 genotypes of *LpSH1* were identified. Four different amino acids sequences were identified for each gene in the corresponding DNA region but no functional changes were identified.

In summary, mutation screening in *LpSH1* and *LpCKX1* was performed using HRM analysis, and different alleles of both target genes were identified and sequenced. However, after sequencing analysis, no useful mutations were detected in the perennial ryegrass population.

6 Chapter 6: Final discussion

The primary aim of this project was to identify key genes involved in seed shattering in perennial ryegrass and then to screen for useful mutations in the target gene of interest from a perennial ryegrass EcoTILLING population. After the project was well underway it was decided to target mutations in *LpCKX1* as well.

6.1 A genetic model for seed shattering in perennial ryegrass

All candidate seed shattering genes were collated from previous studies on crops (Li et al., 2006; Lin et al., 2012; Zhou et al., 2012), and their homologues identified in perennial ryegrass. Phylogenetic analysis showed that all candidate perennial ryegrass genes shared high sequence similarity with orthologues from the Poaceae. Initially, two putative *LpqSH1* DNA sequences were determined from the ryegrass transcriptome database which clustered with *OsqSH1*, suggesting that the two DNAs in perennial ryegrass were homologous to *OsqSH1*. Subsequently, Yoon et al. (2014) identified *OsSH5* as a homologue of *OsqSH1*. After a second phylogenetic analysis, the two DNA sequences clustered with *OsqSH1* and *OsSH5* respectively (Konishi et al., 2006; Yoon et al., 2014). This suggests that phylogenetic analysis is a powerful method to identify homologous genes (Momotaz et al., 2004). Based on this approach, homologous genes can be identified from plant species where little is known about their genomic information, such as for perennial ryegrass. Another group is applying a similar strategy to the one used in this project in an attempt to identify mutations in *qSH1* in an orphan grain crop in Africa, *Digitaria exilis* (fonio), where there are concerns in term of its seed shattering (Patterson et al., 2016).

Two genetic models for seed shattering have been described by Zhou et al. (2012) and Yoon et al. (2017) (Figures 1.3 and 1.4), where the relationships between *qSH1*, *SH5*, *SH4* and *SHAT1* in rice are outlined. Based on the gene expression study described in Chapter 3, it is likely that *LpqSH1* and *LpSH5* play roles upstream of *LpSH4* and *LpSHAT1*, because the first two genes had high levels of expression at the early stage of spike development, and then the expression of the two latter genes increased, corresponding to the timing of expression in rice (Zhou et al., 2012; Yoon et al., 2014; Yoon et al., 2017). In addition, the expression of *LpSH4* was increasing as the seed matured, which is in agreement with that of *SH4* in rice which has a role in degrading the cell wall during the activation of the abscission process (Li et al., 2006). More recently, OSH15, a KNOX protein, was shown to interact independently with

qSH1 and SH5 to induce seed shattering in rice by enhancing both abscission layer differentiation and by inhibiting lignin biosynthesis genes (Yoon et al., 2017). As *OSH15* was identified last year, the homologue of *OSH15* in ryegrass was not included in this study. However, both *LpqSH1* and *LpSH5* were identified as expressing during the early spikelet stage, suggesting that a homologue of *OSH15* might also exist in perennial ryegrass and has a similar role in the abscission process.

In addition to the genes discussed above, *LpSH1* was identified as a key gene in the ryegrass abscission process. *SvSH1* was first identified in sorghum. No abscission layer was observed in the *sh1* mutants, suggesting that *SvSH1* had a role in AL formation (Lin et al., 2012). In addition, the orthologues of *SH1* in rice (*OsSH1*) and in maize (*ZmSH1-1* and *ZmSH1-5.1*+*ZmSH1-5.2*) were identified from conserved collinearity of genomic regions (Lin et al., 2012), suggesting that *SH1* for seed shattering was under parallel selection during the domestication process in crops (Lin et al., 2012). However, Lin et al. (2012) did not perform an expression analysis of *SH1* at different stages of AL development in either sorghum or rice. In this study, the expression of *LpSH1* was shown to increase after anthesis, and the relative expression level was remarkably greater than that of other candidate genes during the seed development stage, suggesting that *LpSH1* probably has a role in activating the abscission process. Tang et al. (2013) also demonstrated that *SbSH1* might maintain the expression of some genes relating to lignin deposition and to play important roles in the abscission process itself. Taken together, it is suggested that *LpSH1* could have a role in abscission layer formation and in the activation of the abscission process in perennial ryegrass.

Three genes, *LpLG1*, *LpQ*, and *LpWRKY*, were ruled out from the candidate gene list. While *LpLG1* has an increased expression for the duration of spike/spikelet development before heading, this is similar to *OsLG1*, a homologue in rice (Ishii et al., 2013). *OsLG1* was identified as playing a role in rice panicle shape, impacting on both seed shattering and pollination (Ishii et al., 2013). *LpLG1* may have a similar role in ryegrass panicle structure, and as a consequence impact both seed shattering and pollination behaviours. If the panicle structure of ryegrass was changed to a spreading panicle, the pollination rate could potentially be increased.

Based on the discussion above, a genetic model for seed shattering in perennial ryegrass is proposed (Figure 6.1). In this diagram, the interaction of LpqSH1 with the KNOX protein

occurs upstream of *LpSH4* and *LpSHAT1*. *LpSH4* and *LpSHAT1* induce the differentiation of the AL in ryegrass. Meanwhile, *LpSH1* could also positively maintain AL formation. It is likely that *LpSH1* continues to have a role during the initiation of abscission, probably by regulating the genes related to lignin deposition (Tang et al., 2013). According to Elgersma et al. (1988), the cell walls of the abscission layer were not lignified in perennial ryegrass. Similar situations have also been observed in rice and wild rye (*Elymus sibiticus*), leading to seed shattering in both species (Yoon et al., 2017; Zhao et al., 2017). Zhao et al. (2017) indicated that a reduced content of lignin in the AL in a wild *Elymus sibiticus* accession, XH09, led to a higher degree of seed shattering, compared with another wild accession, ZhN03, with a low seed shattering degree. Based on these findings, it is suggested that a low lignin deposition in the AL in perennial ryegrass leads to seed shattering. Therefore, in perennial ryegrass, SH5 interacting with a KNOX protein, could repress lignin biosynthesis and thus inhibit lignin deposition, leading to seed shattering. Additionally, *LpSH5* maintains *LpSH4* expression leading to degradation of the cell wall to promote abscission. Interestingly, in the African domesticated rice, the *SH4* mutants also have a smaller seed size, compared with the *SH4* mutant in *Oryza sativa* indica.

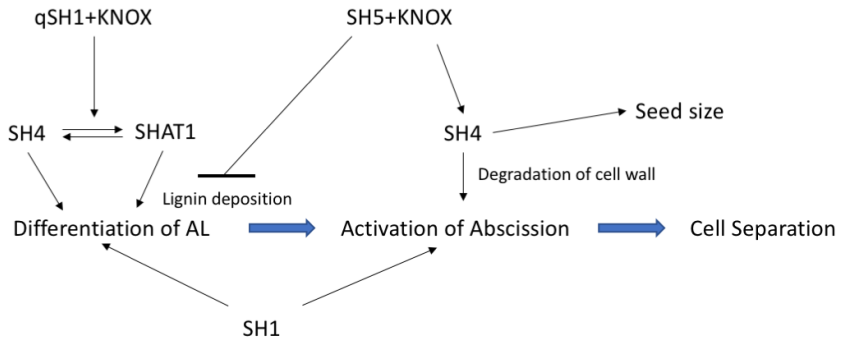


Figure 6.1 Genetic model for seed shattering in perennial ryegrass.

6.2 Cytokinin oxidase/dehydrogenase

For seed size, *CKX1* was selected based on the previous studies in rice and wheat (Ashikari et al., 2005a; Song et al., 2012), and the putative *LpCKX1* was identified from the transcriptome database. The full length of *LpCKX1* was then sequenced.

6.3 Mutation screening

After identification of the key genes and development of the genetic model, this project moved to the second aim of the thesis which was to screen for mutations of *LpSH1* and *LpCKX1* using EcoTILLING strategy. As introduced in Chapters 1 and 4, the drawback of the TILLING approach is that it is time-consuming. In this project, a rapid and cost effective protocol for genomic DNA isolation using regenerated silica columns was developed and applied. This protocol markedly reduced the costs and time of DNA extraction. For example, in this project about 600 genomic DNA samples were isolated from the perennial ryegrass EcoTILLING population in only about two weeks. This research is published (Fu et al., 2017).

In addition to DNA extraction, a method to detect CEL I digestion products was developed. The classic TILLING process uses an electrophoresis platform, such as agarose gel, capillary, or LiCOR electrophoresis, to detect the CEL I digestion products. In the thesis, a method was developed in which the CEL I digestion products were detected using HRM analysis. In addition, the CEL I+HRM method was shown to be able to detect PCR products from pooled DNA samples at 1 in 5 copies. These results suggested that the CEL I+HRM method has potential application for mutation detection. The reliability of this method was tested through mutation detection in exon 1 of *LpCKX1*. Unfortunately, due to a high degree of genetic diversity in perennial ryegrass and the high GC-content of the *LpCKX1* gene, mutations were not detected using this method. Consequently, based on these results, this new method was not recommended to be used for mutation screening in out-crossing species, such as perennial ryegrass.

As the CEL I+HRM method was not suitable for this project, HRM analysis was used to screen for mutations. A homemade master mix was selected and optimised from a variety of DNA polymerases by using two pairs of plasmid, GC and allele4/8. From the optimisation results, the homozygotes of the GC plasmids (GG vs CC) were not discriminated by any of the homemade master mixes or the commercial HRM kits, because the melting temperature of the homozygotes of the supplied GC plasmid differed by less than 0.5ºC. Also, the nearby bases of the SNP site in the GC plasmid are the same, which means that the tiny difference between this plasmid pair cannot be detected by the Rotor-Gene Q. However, for allele4/8, the majority of homemade HRM master mixes could detect the difference within

homozygotes and heterozygotes (4/4, 8/8, 4/8). One homemade HRM master mix with the AmpliTaq DNA polymerase was selected for mutation screening of both target genes in Chapter 5.

Based on the identification of the key genes related to seed shattering and seed size and the optimised HRM analysis for the mutation screening, mutations in both target genes were screened. Twenty-one genotypes of *LpSH1* were identified from the EcoTILLING population. Based on the *LpSH1* HRM analysis results, the cultivars supplied from PGG Wrightson harboured three different genotypes which did not exist in other cultivars from New Zealand. This suggests that it is necessary to source cultivars of widely differing origin to provide novel genotypes for identification of novel mutations.

However, for *LpCKX1*, many melting curve patterns were identified from the HRM analysis, indicating the existence of multiple alleles and genotypes in *LpCKX1*, corresponding to the high degree of genetic diversity shown in perennial ryegrass. In contrast, *LpSH1* presented few melting curve patterns, suggesting that this gene has few genotypes and alleles. It is possible that *LpSH1* is more conserved and under strong selection pressure. Seed shattering is under opposing selections in crops and wild relative. Non-shattering was selected during the domestication by humans and shattering is selected by wild species for seed dispersal (Tang et al., 2013). Of significance also is that *LpSH1* belongs to the YABBY family (Lin et al., 2012), and other YABBY genes have been identified to play roles in a variety of functions in plants, such as *fasciated* in tomato to control locule number (Cong et al., 2008) and *OsYAB1* in rice maintain stamens and carpels (Jang et al., 2004). These examples support the suggestion that *LpSH1* is an essential functional gene in perennial ryegrass. Even though *LpCKX1* presented many melting curve patterns and alleles, the translated amino acid sequences from sequenced samples are conserved, so the evidence of many melting curve patterns does not indicate that *LpCKX1* has an insignificant role in seed size determination in perennial ryegrass. In addition, *LpCKX1* belongs to the *CKX* gene family, so it is possible that other *CKX* gene family members have a similar role to *CKX1* in determining seed phenotype. For example, *CKX3* along with other *CKX* genes controls the size of reproductive shoot apical meristem in arabidopsis in a partially redundant fashion (Bartrina et al., 2011), and multiple *TaCKXs* expressed during floret and post anthesis seed development in wheat (Song et al., 2012).

6.4 Future study

From the gene expression study, *LpSH1* is likely to be a key gene in the abscission process in perennial ryegrass. However, evidence from the expression studies needs to be supported by functional studies of *LpSH1*. As mentioned in Chapter 3, transgenic experiments with *LpSH1* in perennial ryegrass and rice have been implemented. Overexpression-*LpSH1* and RNAi-*LpSH1* vectors have been constructed. The strategy is to over-express *LpSH1* in non-shattering rice (Nipponbare) and to down-regulate *LpSH1* in shattering perennial ryegrass (cv. Nui) using *Agrobacterium*-mediated transformation. If the shattering trait is altered in the transformed plant, it will provide strong support that *LpSH1* is a key gene in the abscission process.

So far, transformed ryegrass with down-regulation of *LpSH1* has not been regenerated. A reason for this may be that the active site of the YABBY domain in *LpSH1* was ligated into the RNAi vector. As the YABBY domain is conserved and has a role in polarity regulation (Toriba et al., 2007), all YABBY family genes may have been down regulated, and affected plant regeneration. It is also possible that the ryegrass cultivar Nui was not suitable for transformation.

Additionally, the overexpression *LpSH1* vector was transformed into a non-shattering rice (Nipponbare), but the experiment failed due to contamination. Recently, the overexpression vector was again transformed into rice using the *Agrobacterium*-mediated transformation method, in collaboration with a biotechnology company in China. This experiment is still on-going. If the overexpressed *LpSH1* in rice modifies the shattering trait, the relationship of *SH1* with other genes, such as *LpqSH1* and *LpSH5* in relation to seed shattering could be revealed.

While Elgersma et al. (1988) showed lignification was not observed in the cell walls of the abscission layer in perennial ryegrass, this histological study needs to be repeated. Histological analysis of the abscission layer in perennial ryegrass should be carried out in two cultivars: cv. Nui with easy-shattering and cv. Med line 1 with moderate-shattering. Based on lignin staining in these two cultivars, the role of lignin deposition can be confirmed in ryegrass, and then compared with other plant species, such as rice (Yoon et al., 2014; Yoon et al., 2017), sorghum (Lin et al., 2012; Tang et al., 2013), and wild rye (*Elymus sibiticus*) (Zhao et al., 2017).

In this study, useful mutations were not identified in either *LpSH1* or *LpCKX1*, but mutations were detected in the perennial ryegrass EcoTILLING population. Therefore, the strategy adopted in this project was appropriate. However, the screened regions in this work only represented a small portion of the coding region for both genes, and across very few genotypes. Functional mutations could be identified by screening other parts of the coding sequences and/or through extending the genotypes to be screened. Also, artificial mutagenesis by EMS, or genome editing using CRISPR/Cas9, could be used to obtain useful mutations of *LpSH1* or *LpCKX1*.

EMS mutagenesis causes random mutations and could generate many different phenotypes with mutations in functional genes. Mutations in alleles beneficial to the industry could then be detected by HRM analysis or CEL I+HRM method. However, mutation screening will still be limited in ryegrass due to its high heterozygosity and self-incompatibility. This means that many alleles with natural nonsense mutations could potentially inhibit the detection of rare useful mutations (Till et al., 2006). Therefore, even in a mutagenised ryegrass population, the rare induced mutations would still be hard to identify by using HRM analysis.

The TILLING strategy was developed mainly to improve the mutation discovery throughput (Wang et al., 2012). In recent years, with the reduction of costs in Next Generation Sequencing (NGS), NGS has been used to screen mutations in TILLING populations (Rigola et al., 2009; Tsai et al., 2011). TILLING by NGS is high-throughput, as the pooled samples can be as high as 40- to 50-fold for some instruments, and at a reasonable cost (Weil, 2009). A high-throughput method with NGS was identified, based on massive parallel sequencing of the target genes (Rigola et al., 2009). Rigola et al. (2009) indicated that two mutations in the *elF4E* gene were identified among more than 3000 M2 families from an EMS-induced tomato population using the Roche 454 technology and a multi-dimensional pooling strategy. In addition, Pembleton et al. (2016) applied SNP genotyping to differentiate ryegrass species and cultivars, using a method similar to that described by Rigola et al. (2009). In their study, PCR products were targeted with identification tags, allowing accelerated throughput and about six-fold reduction in cost.

NGS has the potential to be used in ryegrass EcoTILLING or TILLING populations in the future. For example, considerable variation was detected within the coding region of *LpCKX1*, and some variations may have contained functional changes resulting in useful phenotypes. However, in this project it was not possible to sequence all the *LpCKX1* alleles,

so NGS with a multi-dimensional pooling strategy would be of value for detecting mutations in *LpCKX1*, or in other genes and species, and combining this with HRM analysis to identify mutations from the pooled genotypes.

Compared to the TILLING approach, the advantage of genome editing is to change the DNA sequence in the target gene. Targeted mutations could be introduced using the genome editing, CRISPR/Cas9 system. However, the location of the mutation point in the gene is significant, as different mutations in the same gene can lead to different phenotypes. For example, mutations at different point in the shattering gene, *SH4*, lead to different phenotypic changes, such as changes in the differentiation of the abscission layer, the degradation of the cell wall and seed size (Li et al., 2006; Lin et al., 2007; Wu et al., 2017). This suggests that the site of mutation in the gene could directly affect the phenotype, and this is important for subsequent breeding. CRISPR-Cas9 has been applied to introduce targeted mutations in different plants, such as *TaMLO* and *TaLOX2* in wheat (Shan et al., 2014; Wang et al., 2014) and *OsCKX2* and *OsIPA1* in rice (Li et al., 2016), *DsRED2* in sorghum (Jiang et al., 2013), and *ZmIPK* in maize (Liang et al., 2014).

For *LpSH1*, the YABBY domain has been identified as the key region for gene function, so it is a potential target for genome editing using CRISPR-CAS9 tools in the near future. However, genome editing utilises an initial transgenic step, which is unacceptable in some jurisdictions, including New Zealand, and as such the program will be subjected to the regulations and restrictions that apply to transgenic plants (Royal society of New Zealand, gene editing evidence update, ISBN:978-1-877317-26-2, October 2016). Consequently, TILLING is still an applicable strategy.

References

Acevedo-Garcia, J., Spencer, D., Thieron, H., Reinstadler, A., Hammond-Kosack, K., Phillips, A. L., & Panstruga, R. (2017). mlo-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant Biotechnology Journal, 15*(3), 367-378. doi:10.1111/pbi.12631

Allen, G., Flores-Vergara, M., Krasynanski, S., Kumar, S., & Thompson, W. (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nature Protocols, 1*(5), 2320-2325.

Altpeter, F., Springer, N. M., Bartley, L. E., Blechl, A. E., Brutnell, T. P., Citovsky, V., Conrad, L. J., Gelvin, S. B., Jackson, D. P., Kausch, A. P., Lemaux, P. G., Medford, J. I., Orozco-Cardenas, M. L., Tricoli, D. M., Van Eck, J., Voytas, D. F., Walbot, V., Wang, K., Zhang, Z. Y. J., & Stewart, C. N. (2016). Advancing crop transformation in the era of genome editing. *Plant Cell, 28*(7), 1510-1520. doi:10.1105/tpc.16.00196

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E. R., Qian, Q., Kitano, H., & Matsuoka, M. (2005a). Cytokinin oxidase regulates rice grain production. *Science, 309*. doi:10.1126/science.1113373

Ashikari, M., Sakakibara, H., Lin, S. Y., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E. R., Qian, Q., Kitano, H., & Matsuoka, M. (2005b). Cytokinin oxidase regulates rice grain production. *Science, 309*(5735), 741-745. doi:10.1126/science.1113373

Bajaj, S., Ran, Y. D., Phillips, J., Kularajathevan, G., Pal, S., Cohen, D., Elborough, K., & Puthigae, S. (2006). A high throughput *Agrobacterium tumefaciens*-mediated transformation method for functional genomics of perennial ryegrass (*Lolium perenne* L.). *Plant Cell Reports, 25*(7), 651-659. doi:10.1007/s00299-005-0099-9

Bartrina, I., Otto, E., Strnad, M., Werner, T., & Schmuelling, T. (2011). Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*. *Plant Cell, 23*(1), 69-80. doi:10.1105/tpc.110.079079

Birnboim, H. C. (1988). A Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Current Contents/Life Sciences,* (45), 12-12.

Bleecker, A. B., & Patterson, S. E. (1997). Last exit: Senescence, abscission, and meristem arrest in Arabidopsis. *Plant Cell, 9*(7), 1169-1179. doi:10.1105/tpc.9.7.1169

Blomstedt, C. K., Gleadow, R. M., O'Donnell, N., Naur, P., Jensen, K., Laursen, T., Olsen, C. E., Stuart, P., Hamill, J. D., Moller, B. L., & Neale, A. D. (2012). A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor* L. Moench resulting in acyanogenic forage production. *Plant Biotechnology Journal, 10*(1), 54-66. doi:10.1111/j.1467-7652.2011.00646.x

Bowman, J. L. (2000). The *YABBY* gene family and abaxial cell fate. *Current Opinion in Plant Biology, 3*(1), 17-22. doi:10.1016/s1369-5266(99)00035-7

Brown, K. M. (1997). Ethylene and abscission. *Physiologia Plantarum, 100*(3), 567-576. doi:10.1034/j.1399-3054.1997.1000319.x

Brugière, N., Humbert, S., Rizzo, N., Bohn, J., & Habben, J. E. (2008). A member of the maize isopentenyl transferase gene family, *Zea mays* isopentenyl transferase 2 (*ZmIPT2*), encodes a cytokinin biosynthetic enzyme expressed during kernel development. *Plant Molecular Biology, 67*(3), 215-229. doi:10.1007/s11103-008-9312-x

Brugière, N., Shuping, J., Hanke, S., Zinselmeier, C., Roessler, J. A., Niu, X., Jones, R. J., & Habben, J. E. (2003). Cytokinin oxidase gene expression in maize is localized to the vasculature, and is induced by cytokinins, abscisic acid, and abiotic stress. *Plant Physiology, 132*. doi:10.1104/pp.102.017707

Bush, S. M., & Krysan, P. J. (2010). iTILLING: A personalized approach to the identification of induced mutations in Arabidopsis. *Plant Physiology, 154*(1), 25-35. doi:10.1104/pp.110.159897

Cai, S. Q., & Lashbrook, C. C. (2008). Stamen abscission zone transcriptome profiling reveals new candidates for abscission control: Enhanced retention of floral organs in transgenic plants overexpressing Arabidopsis *ZINC FINGER PROTEIN2*. *Plant Physiology, 146*(3), 1305-1321. doi:10.1104/pp.107.110908

Caldwell, D. G., McCallum, N., Shaw, P., Muehlbauer, G. J., Marshall, D. F., & Waugh, R. (2004). A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *The Plant Journal, 40*(1), 143-150. doi:10.1111/j.1365-313X.2004.02190.x

Charlesworth, D. (1989). Why do plants produce so many more ovules than seeds. *Nature, 338*, 21-22.

Chen, L., Hao, L. G., Parry, M. A. J., Phillips, A. L., & Hu, Y. G. (2014). Progress in TILLING as a tool for functional genomics and improvement of crops. *Journal of Integrative Plant Biology, 56*(5), 425-443. doi:10.1111/jipb.12192

Chen, L., Huang, L., Min, D., Phillips, A., Wang, S., Madgwick, P. J., Parry, M. A. J., & Hu, Y.-G. (2012). Development and characterization of a new TILLING population of common bread wheat (*Triticum aestivum* L.). *PLOS One, 7*(7). doi:10.1371/journal.pone.0041570

Colbert, T., Till, B. J., Tompa, R., Reynolds, S., Steine, M. N., Yeung, A. T., McCallum, C. M., Comai, L., & Henikoff, S. (2001). High-throughput screening for induced point mutations. *Plant Physiology, 126*(2), 480-484. doi:10.1104/pp.126.2.480

Comai, L., Young, K., Till, B. J., Reynolds, S. H., Greene, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., Burtner, C., & Odden, A. R. (2004). Efficient discovery of DNA polymorphisms in natural populations by EcoTILLING. *The Plant Journal, 37*(5), 778-786.

Comai, L., Young, K., Till, B. J., Reynolds, S. H., Greene, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., Burtner, C., Odden, A. R., & Henikoff, S. (2004). Efficient

discovery of DNA polymorphisms in natural populations by EcoTILLING. *The Plant Journal, 37*(5), 778-786. doi:10.1111/j.1365-313X.2003.01999.x

Cong, B., Barrero, L. S., & Tanksley, S. D. (2008). Regulatory change in *YABBY*-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nature Genetics, 40*(6), 800-804. doi:10.1038/ng.144

Cooper, J. P. (1951). Studies on Growth and Development in Lolium: II. Pattern of Bud Development of the Shoot Apex and its Ecological Significance. *Journal of Ecology, 39*(2), 228-270. doi:10.2307/2257911

Cornish, M., Hayward, M., & Lawrence, M. (1979). Self-incompatibility in ryegrass. *Heredity*, 43(1), 129-136.

Deavours, B. E., & Dixon, R. A. (2005). Metabolic engineering of isoflavonoid biosynthesis in alfalfa. *Plant Physiology, 138*(4), 2245-2259.

Dhandapani, P., Song, J., Novak, O., & Jameson, P. E. (2016). Infection by *Rhodococcus fascians* maintains cotyledons as a sink tissue for the pathogen. *Annals of Botany*, mcw202.

Dinneny, J. R., & Yanofsky, M. F. (2005). Drawing lines and borders: how the dehiscent fruit of Arabidopsis is patterned. *Bioessays, 27*(1), 42-49. doi:10.1002/bies.20165

Distefano, G., La Malfa, S., Gentile, A., & Wu, S.-B. (2013). EST-SNP genotyping of citrus species using high-resolution melting curve analysis. *Tree Genetics & Genomes, 9*(5), 1271-1281.

Doebley, J. (2006). Plant science - Unfallen grains: How ancient farmers turned weeds into crops. *Science, 312*(5778), 1318-1319. doi:10.1126/science.1128836

Dong, C., Dalton-Morgan, J., Vincent, K., & Sharp, P. (2009a). A modified TILLING method for wheat breeding. *Plant Genome, 2*(1), 39-47. doi:10.3835/plantgenome2008.10.0012

Dong, C., Vincent, K., & Sharp, P. (2009b). Simultaneous mutation detection of three homoeologous genes in wheat by High Resolution Melting analysis and Mutation Surveyor®. *BMC Plant Biology, 9*(1), 143.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem bull, 19*, 11-15.

Elgersma, A. (1990). Seed yield related to crop development and to yield components in nine cultivars of perennial ryegrass (*Lolium perenne* L.). *Euphytica, 49*(2), 141-154.

Elgersma, A., Leeuwangh, J., & Wilms, H. J. (1988). Abscission and seed shattering in perennial ryegrass (*Lolium perenne* L.). *Euphytica, 39*(3), 51-57.

Estornell, L. H., Agusti, J., Merelo, P., Talon, M., & Tadeo, F. R. (2013). Elucidating mechanisms underlying organ abscission. *Plant Science, 199*, 48-60. doi:10.1016/j.plantsci.2012.10.008

Feng, D. S., Wang, H. G., Zhang, X. S., Kong, L. R., Tian, J. C., & Li, X. F. (2008). Using an inverse PCR method to clone the wheat cytokinin oxidase/dehydrogenase gene *TaCKX1*. *Plant Mol Biol Rep, 26*. doi:10.1007/s11105-008-0033-8

Fu, D. L., Uauy, C., Blechl, A., & Dubcovsky, J. (2007). RNA interference for wheat functional gene analysis. *Transgenic Research, 16*(6), 689-701. doi:10.1007/s11248-007-9150-7

Fu, Z. Y., Song, J. C., & Jameson, P. E. (2017). A rapid and cost effective protocol for plant genomic DNA isolation using regenerated silica columns in combination with CTAB extraction. *Journal of Integrative Agriculture, 16*(8), 1682-1688. doi:10.1016/s2095-3119(16)61534-4

Fukuta, Y., & Yagi, T. (1998). Mapping of a shattering resistance gene in a mutant line SR-5 induced from an indica rice variety, Nan-jing 11. *Japanese Journal of Breeding, 48*(4), 345-348.

Galli, M., & Gallavotti, A. (2016). Expanding the regulatory network for meristem size in plants. *Trends in Genetics, 32*(6), 372-383.

Garritano, S., Gemignani, F., Voegele, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De Silva, D., Lesueur, F., Landi, S., & Tavtigian, S. V. (2009). Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. *BMC Genetics, 10*. doi:10.1186/1471-2156-10-5

Gundry, C. N., Vandersteen, J. G., Reed, G. H., Pryor, R. J., Chen, J., & Wittwer, C. T. (2003). Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clinical Chemistry, 49*(3), 396-406.

Guo, Q. Q., Turnbull, M. H., Song, J. C., Roche, J., Novak, O., Spath, J., Jameson, P. E., & Love, J. (2017). Depletion of carbohydrate reserves limits nitrate uptake during early regrowth in *Lolium perenne* L. *Journal of Experimental Botany, 68*(7), 1569-1583. doi:10.1093/jxb/erx056

Guo, Y. F., & Gan, S. S. (2014). Translational researches on leaf senescence for enhancing plant productivity and quality. *Journal of Experimental Botany, 65*(14), 3901-3913. doi:10.1093/jxb/eru248

Hampton, J., Rolston, M., Pyke, N., & Green, W. (2012). Ensuring the long term viability of the New Zealand seed industry. *Agronomy New Zealand, 42*, 129-140.

Harlan, J. R. (1992). *Crops & man*: American Society of Agronomy.

He, Z., Zhai, W., Wen, H., Tang, T., Wang, Y., Lu, X., Greenberg, A. J., Hudson, R. R., Wu, C.-I., & Shi, S. (2011). Two evolutionary histories in the genome of rice: the roles of domestication genes. *PLOS Genetics, 7*(6). doi:10.1371/journal.pgen.1002100

Hirose, N., Takei, K., Kuroha, T., Kamada-Nobusada, T., Hayashi, H., & Sakakibara, H. (2008). Regulation of cytokinin biosynthesis, compartmentalization and translocation. *Journal of Experimental Botany, 59*(1), 75-83. doi:10.1093/jxb/erm157

Hothorn, M., Dabi, T., & Chory, J. (2011). Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4. *Nature Chemical Biology, 7*(11), 766-768. doi:10.1038/nchembio.667

Ishii, T., Numaguchi, K., Miura, K., Yoshida, K., Thanh, P. T., Htun, T. M., Yamasaki, M., Komeda, N., Matsumoto, T., Terauchi, R., Ishikawa, R., & Ashikari, M. (2013). *OsLG1* regulates a closed panicle trait in domesticated rice. *Nature Genetics, 45*(4), 462-465. doi:10.1038/ng.2567

Ito, M., Sato, Y., & Matsuoka, M. (2002). Involvement of homeobox genes in early body plan of monocot. *International Review of Cytology - a Survey of Cell Biology, Vol 218, 218*, 1-35.

Jameson, P. E., Dhandapani, P., Novak, O., & Song, J. (2016). Cytokinins and expression of *SWEET, SUT, CWINV* and *AAP* genes increase as pea seeds germinate. *International Journal of Molecular Sciences, 17*(12), 2013.

Jameson, P. E., McWha, J. A., & Wright, G. J. (1982). Cytokinins and changes in their activity during the development of grains of wheat (*Triticum aestivum* L.). *J Plant Physiol, 106*.

Jameson, P. E., & Song, J. C. (2016). Cytokinin: a key driver of seed yield. *Journal of Experimental Botany, 67*(3), 593-606. doi:10.1093/jxb/erv461

Jang, S., Hur, J., Kim, S. J., Han, M. J., Kim, S. R., & An, G. (2004). Ectopic expression of *OsYAB1* causes extra stamens and carpels in rice. *Plant Molecular Biology, 56*(1), 133-143. doi:10.1007/s11103-004-2648-y

Ji, H., Kim, S.-R., Kim, Y.-H., Kim, H., Eun, M.-Y., Jin, I.-D., Cha, Y.-S., Yun, D.-W., Ahn, B.-O., Lee, M. C., Lee, G.-S., Yoon, U.-H., Lee, J.-S., Lee, Y.-H., Suh, S.-C., Jiang, W., Yang, J.-I., Jin, P., McCouch, S. R., An, G., & Koh, H.-J. (2010). Inactivation of the CTD phosphatase-like gene *OsCPL1* enhances the development of the abscission layer and seed shattering in rice. *The Plant Journal, 61*(1), 96-106. doi:10.1111/j.1365-313X.2009.04039.x

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research, 41*(20), e188-e188.

Jones, E. S., Mahoney, N. L., Hayward, M. D., Armstead, I. P., Jones, J. G., Humphreys, M. O., King, I. P., Kishida, T., Yamada, T., Balfourier, F., Charmet, G., & Forster, J. W. (2002). An enhanced molecular marker based genetic map of perennial ryegrass (*Lolium perenne*) reveals comparative relationships with other Poaceae genomes. *Genome, 45*(2), 282-295. doi:10.1139/g01-144

Knoll, J. E., Ramos, M. L., Zeng, Y., Holbrook, C. C., Chow, M., Chen, S., Maleki, S., Bhattacharya, A., & Ozias-Akins, P. (2011). TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). *BMC Plant Biology, 11*. doi:10.1186/1471-2229-11-81

Konishi, S., Izawa, T., Lin, S. Y., Ebana, K., Fukuta, Y., Sasaki, T., & Yano, M. (2006). An SNP caused loss of seed shattering during rice domestication. *Science, 312*(5778), 1392-1396. doi:10.1126/science.1126410

Kumar, A. P. K., Boualem, A., Bhattacharya, A., Parikh, S., Desai, N., Zambelli, A., Leon, A., Chatterjee, M., & Bendahmane, A. (2013). SMART - Sunflower Mutant population And Reverse genetic Tool for crop improvement. *BMC Plant Biology, 13*. doi:10.1186/1471-2229-13-38

Lababidi, S., Mejlhede, N., Rasmussen, S. K., Backes, G., Al-Said, W., Baum, M., & Jahoor, A. (2009). Identification of barley mutants in the cultivar 'Lux' at the *Dhn* loci through TILLING. *Plant Breeding, 128*(4), 332-336. doi:10.1111/j.1439-0523.2009.01640.x

Lashbrook, C. C., & Cai, S. (2008). Cell wall remodeling in Arabidopsis stamen abscission zones: Temporal aspects of control inferred from transcriptional profiling. *Plant Signaling & Behavior, 3*(9).

Lemke, L., Rex, M., Zyprian, E., & Töpfer, R. (2011). A simple, inexpensive and environmentally friendly method for high throughput DNA extraction from grapevine (*Vitis* spp.). *Vitis, 50*(1), 7-10.

Letham, D. S. (1963). Zeatin, a factor inducing cell division isolated from *Zea Mays*. *Life Sciences*(8), 569-573.

Li, C. B., Zhou, A. L., & Sang, T. (2006). Rice domestication by reducing shattering. *Science, 311*(5769), 1936-1939. doi:10.1126/science.1123604

Li, J., Nie, X., Tan, J. L. H., & Berger, F. (2013). Integration of epigenetic and genetic controls of seed size by cytokinin in Arabidopsis. *PNAS, 110*(38), 15479-15484. doi:10.1073/PNAS.1305175110

Li, M. R., Li, X. X., Zhou, Z. J., Wu, P. Z., Fang, M. C., Pan, X. P., Lin, Q. P., Luo, W. B., Wu, G. J., & Li, H. Q. (2016). Reassessment of the four yield-related genes *Gn1a*, *DEP1*, *GS3*, and *IPA1* in rice using a CRISPR/Cas9 system. *Frontiers in Plant Science, 7*. doi:10.3389/fpls.2016.00377

Liang, Z., Zhang, K., Chen, K., & Gao, C. (2014). Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *Journal of Genetics and Genomics, 41*(2), 63-68.

Liebman, M., & Dyck, E. (1993). Crop-rotation and intercropping strategies for weed management. *Ecological Applications, 3*(1), 92-122. doi:10.2307/1941795

Liew, M., Pryor, R., Palais, R., Meadows, C., Erali, M., Lyon, E., & Wittwer, C. (2004). Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clinical Chemistry, 50*(7), 1156-1164. doi:10.1373/clinchem.2004.032136

Lin, Z., Griffith, M. E., Li, X., Zhu, Z., Tan, L., Fu, Y., Zhang, W., Wang, X., Xie, D., & Sun, C. (2007). Origin of seed shattering in rice (*Oryza sativa* L.). *Planta, 226*(1), 11-20. doi:10.1007/s00425-006-0460-4

Lin, Z., Li, X., Shannon, L. M., Yeh, C.-T., Wang, M. L., Bai, G., Peng, Z., Li, J., Trick, H. N., & Clemente, T. E. (2012). Parallel domestication of the *Shattering1* genes in cereals. *Nature Genetics, 44*(6), 720-724.

Liu, J., Barker, D., Sulc, R., Jang, J., & Wang, G. (2010). Quantifying the proportion of perennial ryegrass cultivars in intra-species mixtures using simple sequence repeat (SSR) and inter-SSR (ISSR) markers and discriminant analysis. *New Zealand Journal of Agricultural Research, 53*(3), 215-226.

Liu, Z. N., Lv, Y. X., Zhang, M., Liu, Y. P., Kong, L. J., Zou, M. H., Lu, G., Cao, J. S., & Yu, X. L. (2013). Identification, expression, and comparative genomic analysis of the *IPT* and *CKX* gene families in Chinese cabbage (*Brassica rapa* ssp pekinensis). *BMC Genomics, 14*. doi:10.1186/1471-2164-14-594

Lochlainn, S. O., Amoah, S., Graham, N. S., Alamer, K., Rios, J. J., Kurup, S., Stoute, A., Hammond, J. P., Ostergaard, L., King, G. J., White, P. J., & Broadley, M. R. (2011). High Resolution Melt (HRM) analysis is an efficient tool to genotype EMS mutants in complex crop genomes. *Plant Methods, 7*. doi:10.1186/1746-4811-7-43

Lomin, S. N., Krivosheev, D. M., Steklov, M. Y., Arkhipov, D. V., Osolodkin, D. I., Schmulling, T., & Romanov, G. A. (2015). Plant membrane assays with cytokinin receptors underpin the unique role of free cytokinin bases as biologically active ligands. *Journal of Experimental Botany, 66*(7), 1851-1863. doi:10.1093/jxb/eru522

Mader, E., Lukas, B., & Novak, J. (2008). A strategy to setup codominant microsatellite analysis for High-Resolution-Melting-Curve-analysis (HRM). *BMC Genetics, 9*. doi:10.1186/1471-2156-9-69

Manzanares, C., Yates, S., Ruckle, M., Nay, M., & Studer, B. (2016). TILLING in forage grasses for gene discovery and breeding improvement. *New Biotechnology, 33*(5), 594-603. doi:10.1016/j.nbt.2016.02.009

Mao, F., Leung, W.-Y., & Xin, X. (2007). Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC Biotechnology, 7*(1), 76.

Mao, L., Begum, D., Chuang, H. W., Budiman, M. A., Szymkowiak, E. J., Irish, E. E., & Wing, R. A. (2000). *JOINTLESS* is a MADS-box gene controlling tomato flower abscission zone development. *Nature, 406*(6798), 910-913.

McCallum, C. M., Comai, L., Greene, E. A., & Henikoff, S. (2000). Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiology, 123*(2), 439-442. doi:10.1104/pp.123.2.439

McKim, S. M., Stenvik, G.-E., Butenko, M. A., Kristiansen, W., Cho, S. K., Hepworth, S. R., Aalen, R. B., & Haughn, G. W. (2008). The *BLADE-ON-PETIOLE* genes are essential for abscission zone formation in Arabidopsis. *Development, 135*(8), 1537-1546. doi:10.1242/dev.012807

Meir, S., Hunter, D. A., Chen, J. C., Halaly, V., & Reid, M. S. (2006). Molecular changes occurring during acquisition of abscission competence following auxin depletion in *Mirabilis jalapa*. *Plant Physiology, 141*(4), 1604-1616. doi:10.1104/pp.106.079277

Meir, S., Philosoph-Hadas, S., Sundaresan, S., Selvaraj, K. S. V., Burd, S., Ophir, R., Kochanek, B., Reid, M. S., Jiang, C. Z., & Lers, A. (2010). Microarray analysis of the abscission-related transcriptome in the tomato flower abscission zone in response to auxin depletion. *Plant Physiology, 154*(4), 1929-1956. doi:10.1104/pp.110.160697

Miller, C. O., Skoog, F., Vonsaltza, M. H., & Strong, F. M. (1955). Kinetin, a cell division factor from deoxyribonucleic acid. *Journal of the American Chemical Society, 77*(5), 1392-1392. doi:10.1021/ja01610a105

Minoia, S., Petrozza, A., D'Onofrio, O., Piron, F., Mosca, G., Sozio, G., Cellini, F., Bendahmane, A., & Carriero, F. (2010). A new mutant genetic resource for tomato crop improvement by TILLING technology. *BMC Research Notes, 3*, 69-69. doi:10.1186/1756-0500-3-69

Miyawaki, K., Kakimoto, T., & Matsumoto-Kitano, M. (2004). Expression of cytokinin biosynthetic isopentenyltransferase genes in Arabidopsis: tissue specificity and regulation by auxin, cytokinin and nitrate. *The Plant Journal, 37*. doi:10.1046/j.1365-313X.2003.01945.x

Momotaz, A., Forster, J. W., & Yamada, T. (2004). Identification of cultivars and accessions of *Lolium*, *Festuca* and *Festulolium* hybrids through the detection of simple sequence repeat polymorphism. *Plant Breeding, 123*(4), 370-376. doi:10.1111/j.1439-0523.2004.00962.x

Monis, P. T., Giglio, S., & Saint, C. P. (2005). Comparison of SYTO9 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of dye concentration on amplification and DNA melting curve analysis. *Analytical Biochemistry, 340*(1), 24-34.

Morrell, P. L., Buckler, E. S., & Ross-Ibarra, J. (2012). Crop genomics: advances and applications. *Nature Reviews Genetics, 13*(2), 85-96. doi:10.1038/nrg3097

Morris, R. O., Blevins, D. G., Dietrich, J. T., Durley, R. C., Gelvin, S. B., Gray, J., Hommes, N. G., Kaminek, M., Mathews, L. J., Meilan, R., Reinbott, T. M., & Sayavedra-Soto, L. (1993). Cytokinins in plant-pathogenic bacteria and developing cereal grains. *Australian Journal of Plant Physiology, 20*(5). doi:10.1071/pp9930621

Moubayidin, L., Di Mambro, R., & Sabatini, S. (2009). Cytokinin-auxin crosstalk. *Trends in Plant Science, 14*(10), 557-562. doi:10.1016/j.tplants.2009.06.010

MPI. (2016). *How valuable is that plant species? Application of a method for enumerating the contribution of selected plant species to New Zealand's GDP*. Retrieved from https://www.mpi.govt.nz/dmsdocument/14527-how-valuable-is-that-plant-species-application-of-a-method-for-enumerating-the-contribution-of-selected-plant-species-to-new-zealands-gdp

Muller, B., & Sheen, J. (2008). Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis. *Nature, 453*(7198), 1094-U1097. doi:10.1038/nature06943

Muth, J., Hartje, S., Twyman, R. M., Hofferbert, H.-R., Tacke, E., & Pruefer, D. (2008). Precision breeding for novel starch variants in potato. *Plant Biotechnology Journal, 6*(6), 576-584. doi:10.1111/j.1467-7652.2008.00340.x

Nakano, T., & Ito, Y. (2013). Molecular mechanisms controlling plant organ abscission. *Plant Biotechnology, 30*(3), 209-216. doi:10.5511/plantbiotechnology.13.0318a

Nakano, T., Kimbara, J., Fujisawa, M., Kitagawa, M., Ihashi, N., Maeda, H., Kasumi, T., & Ito, Y. (2012). MACROCALYX and JOINTLESS interact in the transcriptional regulation of tomato fruit abscission zone development. *Plant Physiology, 158*(1), 439-450. doi:10.1104/pp.111.183731

O'Keefe, D., Song, J., & Jameson, P. E. (2011). Isopentenyl transferase and cytokinin oxidase/dehydrogenase gene family members are differentially expressed during pod and seed development in rapid-cycling brassica. *Journal of Plant Growth Regulation, 30*. doi:10.1007/s00344-010-9171-y

Ogawa, M., Kay, P., Wilson, S., & Swain, S. M. (2009). ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 are polygalacturonases required for cell separation during reproductive development in Arabidopsis. *Plant Cell, 21*(1), 216-233. doi:10.1105/tpc.108.063768

Ohmori, Y., Abiko, M., Horibata, A., & Hirano, H. Y. (2008). A transposon, Ping, is integrated into intron 4 of the *DROOPING LEAF* gene of rice, weakly reducing its expression and causing a mild drooping leaf phenotype. *Plant and Cell Physiology, 49*(8), 1176-1184. doi:10.1093/pcp/pcn093

Oleykowski, C. A., Mullins, C. R. B., Godwin, A. K., & Yeung, A. T. (1998). Mutation detection using a novel plant endonuclease. *Nucleic Acids Research, 26*(20), 4597-4602. doi:10.1093/nar/26.20.4597

Onishi, K., Takagi, K., Kontani, M., Tanaka, T., & Sano, Y. (2007). Different patterns of genealogical relationships found in the two major QTLs causing reduction of seed shattering during rice domestication. *Genome, 50*(8), 757-766. doi:10.1139/g07-051

Parry, M. A. J., Madgwick, P. J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., Labhilili, M., & Phillips, A. L. (2009). Mutation discovery for crop improvement. *Journal of Experimental Botany, 60*(10), 2817-2825. doi:10.1093/jxb/erp189

Patterson, S. E. (2001). Cutting loose. Abscission and Dehiscence in Arabidopsis. *Plant Physiology, 126*(2), 494-500. doi:10.1104/pp.126.2.494

Patterson, S. E., & Bleecker, A. B. (2004). Ethylene-dependent and -independent processes associated with floral organ abscission in Arabidopsis. *Plant Physiology, 134*(1), 194-203. doi:10.1104/pp.103.028027

Patterson, S. E., Bolivar-Medina, J. L., Falbel, T. G., Hedtcke, J. L., Nevarez-McBride, D., Maule, A. F., & Zalapa, J. E. (2016). Are we on the right track: can our understanding of abscission in model systems promote or derail making improvements in less studied crops? *Frontiers in Plant Science, 6*. doi:10.3389/fpls.2015.01268

Pembleton, L. W., Drayton, M. C., Bain, M., Baillie, R. C., Inch, C., Spangenberg, G. C., Wang, J. P., Forster, J. W., & Cogan, N. O. I. (2016). Targeted genotyping-by-sequencing permits cost-effective identification and discrimination of pasture grass species and cultivars. *Theoretical and Applied Genetics, 129*(5), 991-1005. doi:10.1007/s00122-016-2678-2

Perry, J. A., Wang, T. L., Welham, T. J., Gardner, S., Pike, J. M., Yoshida, S., & Parniske, M. (2003). A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiology, 131*(3), 866-871. doi:10.1104/pp.102.017384

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research, 29*(9), e45-e45.

Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E., & Yanofsky, M. F. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature, 424*(6944), 85-88. doi:10.1038/nature01741

Purugganan, M. D. (2014). An evolutionary genomic tale of two rice species. *Nature Genetics, 46*(9), 931-932. doi:10.1038/ng.3071

Qiu, P., Shandilya, H., D'Alessio, J. M., O'Connor, K., Durocher, J., & Gerard, G. F. (2004). Mutation detection using Surveyor™ nuclease. *BioTechniques, 36*(4), 702-707.

Rapley, R., & Harbron, S. (2005). *Molecular analysis and genome discovery*: John Wiley & Sons.

Reed, G. H., & Wittwer, C. T. (2004). Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. *Clinical Chemistry, 50*(10), 1748-1754. doi:10.1373/clinchem.2003.029751

Reiser, L., Modrusan, Z., Margossian, L., Samach, A., Ohad, N., Haughn, G. W., & Fischer, R. L. (1995). The *BELL1* gene encodes a homeodomain protein involved in pattern-formation in the Arabidopsis ovule primordium. *Cell, 83*(5), 735-742. doi:10.1016/0092-8674(95)90186-8

Rigola, D., van Oeveren, J., Janssen, A., Bonne, A., Schneiders, H., van der Poel, H. J. A., van Orsouw, N. J., Hogers, R. C. J., de Both, M. T. J., & van Eijk, M. J. T. (2009). High-throughput detection of induced mutations and natural variation using KeyPoint™ technology. *PLOS One, 4*(3). doi:10.1371/journal.pone.0004761

Roberts, J. A., Elliott, K. A., & Gonzalez-Carranza, Z. H. (2002). Abscission,dehiscence, and other cell separation processes. *Annual Review of Plant Biology, 53*, 131-158. doi:10.1146/annurev.arplant.53.092701.180236

Roche, J., Love, J., Guo, Q. Q., Song, J. C., Cao, M. S., Fraser, K., Huege, J., Jones, C., Novak, O., Turnbull, M. H., & Jameson, P. E. (2016). Metabolic changes and associated cytokinin signals in response to nitrate assimilation in roots and shoots of *Lolium perenne*. *Physiologia Plantarum, 156*(4), 497-511. doi:10.1111/ppl.12412

Roeder, A. H. K., Ferrandiz, C., & Yanofsky, M. F. (2003). The role of the REPLUMLESS homeodomain protein in patterning the Arabidopsis fruit. *Current Biology, 13*(18), 1630-1635. doi:10.1016/j.cub.2003.08.027

Sakakibara, H. (2006). Cytokinins: activity, biosynthesis, and translocation. *Annual Review of Plant Biology, 57*, 431-449. doi:10.1146/annurev.arplant.57.032905.105231

Shan, Q. W., Wang, Y. P., Li, J., & Gao, C. X. (2014). Genome editing in rice and wheat using the CRISPR/Cas system. *Nature Protocols, 9*(10), 2395-2410. doi:10.1038/nprot.2014.157

Shimizu-Sato, S., Tanaka, M., & Mori, H. (2009). Auxin-cytokinin interactions in the control of shoot branching. *Plant Molecular Biology, 69*(4), 429-435. doi:10.1007/s11103-008-9416-3

Siddappa, N. B., Avinash, A., Venkatramanan, M., & Ranga, U. (2007). Regeneration of commercial nucleic acid extraction columns without the risk of carryover contamination. *BioTechniques, 42*(2), 186.

Simko, I. (2016). High-Resolution DNA melting analysis in plant research. *Trends in Plant Science, 21*(6), 528-537. doi:10.1016/j.tplants.2016.01.004

Simon, U. H., M. D.; Kjaersgaard, B.; Clifford, P. T. P.; Hampton, J. G.; Hill, M. J. (1997). Forage seed production. *Harvest and Postharvest Management of Forage Seed Crops*. Wallingford: CAB International.

Simons, K. J., Fellers, J. P., Trick, H. N., Zhang, Z. C., Tai, Y. S., Gill, B. S., & Faris, J. D. (2006). Molecular characterization of the major wheat domestication gene *Q*. *Genetics, 172*(1), 547-555. doi:10.1534/genetics.105.044727

Slade, A. J., Fuerstenberg, S. I., Loeffler, D., Steine, M. N., & Facciotti, D. (2005). A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nature Biotechnology, 23*. doi:10.1038/nbt1043

Smidansky, E. D., Clancy, M., Meyer, F. D., Lanning, S. P., Blake, N. K., Talbert, L. E., & Giroux, M. J. (2002). Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *PNAS, 99*. doi:10.1073/pnas.022635299

Song, J., Clemens, J., & Jameson, P. E. (2008). Quantitative expression analysis of the *ABC* genes in *Sophora tetraptera*, a woody legume with an unusual sequence of floral organ development. *Journal of Experimental Botany, 59*. doi:10.1093/jxb/erm305

Song, J., Jiang, L., & Jameson, P. E. (2012). Co-ordinate regulation of cytokinin gene family members during flag leaf and reproductive development in wheat. *BMC Plant Biology, 12*(1), 78.

Song, J., Jiang, L., & Jameson, P. E. (2015). Expression patterns of *Brassica napus* genes implicate *IPT* sucrose transporter, cell wall invertase, and amino acid permease gene family members in leaf, flower, silique, and seed development. *Journal of Experimental Botany, 66*(16), 5067-5082. doi:10.1093/jxb/erv133

Spichal, L. (2012). Cytokinins - recent news and views of evolutionally old molecules. *Functional Plant Biology, 39*(4), 267-284. doi:10.1071/fp11276

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G. J., & Ostergaard, L. (2010). A rich TILLING resource for studying gene function in *Brassica rapa*. *BMC Plant Biology, 10*. doi:10.1186/1471-2229-10-62

Sundaresan, V. (2005). Control of seed size in plants. *PNAS, 102*(50), 17887-17888. doi:10.1073/pnas.0509021102

Tang, H. B., Cuevas, H. E., Das, S., Sezen, U. U., Zhou, C. B., Guo, H., Goff, V. H., Ge, Z. X., Clemente, T. E., & Paterson, A. H. (2013). Seed shattering in a wild sorghum is conferred by a locus unrelated to domestication. *PNAS, 110*(39), 15824-15829. doi:10.1073/pnas.1305213110

Taylor, J. E., & Whitelaw, C. A. (2001). Signals in abscission. *New Phytologist, 151*(2), 323-339. doi:10.1046/j.0028-646x.2001.00194.x

Tesniere, C., Torregrosa, L., Pradal, M., Souquet, J.-M., Gilles, C., Dos Santos, K., Chatelet, P., & Gunata, Z. (2006). Effects of genetic manipulation of alcohol dehydrogenase levels on the response to stress and the synthesis of secondary metabolites in grapevine leaves. *Journal of Experimental Botany, 57*(1), 91-99.

Till, B. J., Burtner, C., Comai, L., & Henikoff, S. (2004). Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Research, 32*(8), 2632-2641. doi:10.1093/nar/gkh599

Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S., & Comai, L. (2007). Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biology, 7*. doi:10.1186/1471-2229-7-19

Till, B. J., Reynolds, S. H., Greene, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., Burtner, C., Odden, A. R., Young, K., Taylor, N. E., Henikoff, J. G., Comai, L., & Henikoff, S. (2003). Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Research, 13*(3), 524-530. doi:10.1101/gr.977903

Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C. A., Enns, L. C., Odden, A. R., Greene, E. A., Comai, L., & Henikoff, S. (2004a). Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biology, 4*, doi:10.1186/1471-2229-4-12

Till, B. J., Zerr, T., Comai, L., & Henikoff, S. (2006). A protocol for TILLING and EcoTILLING in plants and animals. *Nature Protocols, 1*(5), 2465-2477. doi:10.1038/nprot.2006.329

Toriba, T., Harada, K., Takamura, A., Nakamura, H., Ichikawa, H., Suzaki, T., & Hirano, H. Y. (2007). Molecular characterization the *YABBY* gene family in *Oryza sativa* and expression analysis of *OsYABBY1*. *Molecular Genetics and Genomics, 277*(5), 457-468. doi:10.1007/s00438-006-0202-0

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K. J., Lieberman, M., Fass, J., Uauy, C., Tran, R. K., Khan, A. A., Filkov, V., Tai, T. H., Dubcovsky, J., &

Comai, L. (2011). Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiology, 156*(3), 1257-1268. doi:10.1104/pp.110.169748

Tucker, M. L., Burke, A., Murphy, C. A., Thai, V. K., & Ehrenfried, M. L. (2007). Gene expression profiles for cell wall-modifying proteins associated with soybean cyst nematode infection, petiole abscission, root tips, flowers, apical buds, and leaves. *Journal of Experimental Botany, 58*(12), 3395-3406. doi:10.1093/jxb/erm188

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R. K., Tsai, H., Berardi, S., Comai, L., & Dubcovsky, J. (2009). A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biology, 9*. doi:10.1186/1471-2229-9-115

Veit, B. (2009). Hormone mediated regulation of the shoot apical meristem. *Plant Molecular Biology, 69*(4), 397-408. doi:10.1007/s11103-008-9396-3

Wang, T. L., Uauy, C., Robson, F., & Till, B. (2012). TILLING *in extremis*. *Plant Biotechnology Journal, 10*(7), 761-772. doi:10.1111/j.1467-7652.2012.00708.x

Wang, Y. P., Cheng, X., Shan, Q. W., Zhang, Y., Liu, J. X., Gao, C. X., & Qiu, J. L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology, 32*(9), 947-951. doi:10.1038/nbt.2969

Weil, C. F. (2009). TILLING in grass species. *Plant Physiology, 149*(1), 158-164. doi:10.1104/pp.108.128785

Werner, T., Holst, K., Pörs, Y., Guivarc'h, A., Mustroph, A., Chriqui, D., Grimm, B., & Schmülling, T. (2008). Cytokinin deficiency causes distinct changes of sink and source parameters in tobacco shoots and roots. *Journal of Experimental Botany, 59*(10), 2659-2672.

Werner, T., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H., & Schmulling, T. (2003). Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell, 15*(11), 2532-2550. doi:10.1105/tpc.014928

Wilde, H. D., Chen, Y., Jiang, P., & Bhattacharya, A. (2012). Targeted mutation breeding of horticultural plants. *Emirates Journal of Food and Agriculture, 24*(1), 31.

Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G., & Pryor, R. J. (2003). High-resolution genotyping by amplicon melting analysis using LCGreen. *Clinical Chemistry, 49*(6), 853-860.

Wu, J. L., Wu, C. J., Lei, C. L., Baraoidan, M., Bordeos, A., Madamba, M. R. S., Ramos-Pamplona, M., Mauleon, R., Portugal, A., Ulat, V. J., Bruskiewich, R., Wang, G. L., Leach, J., Khush, G., & Leung, H. (2005). Chemical- and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics. *Plant Molecular Biology, 59*(1), 85-97. doi:10.1007/s11103-004-5112-0

Wu, W. G., Liu, X. Y., Wang, M. H., Meyer, R. S., Luo, X. J., Ndjiondjop, M. N., Tan, L. B., Zhang, J. W., Wu, J. Z., Cai, H. W., Sun, C. Q., Wang, X. K., Wing, R. A., & Zhu, Z. F. (2017). A single-nucleotide polymorphism causes smaller grain size and loss of

seed shattering during African rice domestication. *Nature Plants, 3*(6). doi:10.1038/nplants.2017.64

Xin, Z., Wang, M. L., Barkley, N. A., Burow, G., Franks, C., Pederson, G., & Burke, J. (2008). Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biology, 8*. doi:10.1186/1471-2229-8-103

Yant, L., Mathieu, J., Dinh, T. T., Ott, F., Lanz, C., Wollmann, H., Chen, X. M., & Schmid, M. (2010). Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor *APETALA2*. *Plant Cell, 22*(7), 2156-2170. doi:10.1105/tpc.110.075606

Yoon, J., Cho, L.-H., Antt, H. W., Koh, H.-J., & An, G. (2017). KNOX protein OSH15 induces grain shattering by repressing lignin biosynthesis genes. *Plant Physiology*, DOI: https://doi.org/10.1104/pp.17.00298

Yoon, J., Cho, L. H., Kim, S. L., Choi, H., Koh, H. J., & An, G. (2014). The BEL1‐type homeobox gene *SH5* induces seed shattering by enhancing abscission‐zone development and inhibiting lignin biosynthesis. *The Plant Journal, 79*(5), 717-728.

Zalewski, W., Galuszka, P., Gasparis, S., Orczyk, W., & Nadolska-Orczyk, A. (2010). Silencing of the *HvCKX1* gene decreases the cytokinin oxidase/dehydrogenase level in barley and leads to higher plant productivity. *Journal of Experimental Botany, 61*. doi:10.1093/jxb/erq052

Zhang, L., Liu, D. M., Wang, D., Zhang, R. Z., Geng, S. F., Wu, L., Li, A. L., & Mao, L. (2013). Over expression of the wheat BEL1-like gene *TaqSH1* affects floral organ abscission in *Arabidopsis thaliana*. *Journal of Plant Biology, 56*(2), 98-105. doi:10.1007/s12374-012-0438-7

Zhang, L., Zhao, Y. L., Gao, L. F., Zhao, G. Y., Zhou, R. H., Zhang, B. S., & Jia, J. Z. (2012). *TaCKX6-D1*, the ortholog of rice *OsCKX2*, is associated with grain weight in hexaploid wheat. *New Phytologist, 195*(3), 574-584. doi:10.1111/j.1469-8137.2012.04194.x

Zhao, X. H., Xie, W. G., Zhang, J. C., Zhang, Z. Y., & Wang, Y. R. (2017). Histological characteristics, cell wall hydrolytic enzymes activity and candidate genes expression associated with seed shattering of *Elymus sibiricus* accessions. *Frontiers in Plant Science, 8*. doi:10.3389/fpls.2017.00606

Zhou, Y., Lu, D., Li, C., Luo, J., Zhu, B.-F., Zhu, J., Shangguan, Y., Wang, Z., Sang, T., Zhou, B., & Han, B. (2012). Genetic control of seed shattering in rice by the APETALA2 transcription factor *SHATTERING ABORTION1*. *Plant Cell, 24*(3), 1034-1048. doi:10.1105/tpc.111.094383

Appendices

Appendix 2.1 Spectrophotometric ratios of purified DNA: DNA obtained from a new column and columns regenerated up to 11 times (Re1-11); three replicates.

Replicate	Number of regeneration	Concentration (ng/µl)	A260	A280	A230	A260/280	A260/230	340 raw
1	New	756.07	15.121	7.125	6.277	2.12	2.41	0.093
	Re1	682.86	13.657	6.573	6.75	2.08	2.02	0.194
	Re2	573.29	11.466	5.714	5.162	2.01	2.22	0.018
	Re3	520.93	10.419	5.183	4.647	2.01	2.24	0.055
	Re4	497.85	9.957	4.893	4.192	2.04	2.38	0.057
	Re5	376.16	7.523	3.691	3.16	2.04	2.38	0.093
	Re6	358.42	7.168	3.48	3.02	2.06	2.37	0.098
	Re7	400	3.998	1.964	1.673	2.04	2.39	0.039
	Re8	304.97	6.099	2.972	2.554	2.05	2.39	0.046
	Re9	267.86	5.357	2.604	2.253	2.06	2.38	0.072
	Re10	263.53	5.271	2.56	2.236	2.06	2.36	0.15
	Re11	294.21	5.884	2.856	2.467	2.06	2.39	0.051
2	New	763.89	15.278	7.396	6.681	2.07	2.29	2.108
	Re1	727.98	14.56	7.018	6.099	2.07	2.39	0.249
	Re2	703.97	14.079	6.796	5.922	2.07	2.38	0.144
	Re3	496.62	9.932	4.886	4.206	2.03	2.36	0.056
	Re4	509.51	10.19	4.992	4.3	2.04	2.37	0.048
	Re5	347.84	6.957	3.368	2.975	2.07	2.34	0.048
	Re6	331.87	6.637	3.242	2.77	2.05	2.4	0.035
	Re7	279.69	5.594	2.704	2.343	2.07	2.39	0.02
	Re8	354.27	7.085	3.492	2.957	2.03	2.4	0.032
	Re9	279.73	5.595	2.719	2.344	2.06	2.39	0.034
	Re10	374.72	7.494	3.628	3.163	2.07	2.37	0.054
	Re11	283.6	5.672	2.731	2.367	2.08	2.4	0.04
3	New	963.98	19.28	9.098	8.188	2.12	2.35	0.265
	Re1	942.04	18.841	9.049	7.918	2.08	2.38	0.556
	Re2	960.75	19.215	9.405	8.212	2.04	2.34	0.17
	Re3	791.91	15.838	7.587	7.165	2.09	2.21	0.298
	Re4	567.19	11.344	5.593	4.856	2.03	2.34	0.059
	Re5	559.86	11.197	5.498	5.947	2.04	1.88	0.041
	Re6	462.79	9.256	4.57	4.54	2.03	2.04	0.036
	Re7	421.92	8.438	4.134	3.755	2.04	2.25	0.041
	Re8	372.02	7.44	3.581	3.244	2.08	2.29	0.061
	Re9	384.86	7.697	3.741	3.378	2.06	2.28	0.052
	Re10	317.3	6.346	3.077	2.768	2.06	2.29	0.056
	Re11	439.31	8.786	4.311	4.624	2.04	1.9	0.062

Appendix 3.1 Primer sequence for *LpSH1* introns and untranslated regions.

Name	Sequence	Amplicon region
LpSH1F3	CCACATGCACGCACACTTGCACTAG	5' UTR
LpSH1F4	ACTTGCACTAGCAATTTGCTCCTATGTGTC	
LpSH1F5	GCTAGTGGTCCTTCGTCTCTTCTTCTACTTC	
LpSH1F6	CTTCATCTTCTTTCCTTCATCTTCCTCATCCTC	
LpSH1R3	GCAGTGCACGTAGCACACATGCTC	
LpSH1R4	GCGAGAATTGTGTTGCAGAAGTTGCAG	
LpSH1F7	CCGGCGGAGCATGTGTGCTAC	Intron 1
LpSH1F8	TCTGCAACACAATTCTCGCGGTCAG	
LpSH1R5	CACCGACAGCAAGCTAGTGCAGTG	
LpSH1R6	GTGGTGGTGACTGGATTAAACCTCTCAAG	
LpSH1F9	TGCTGAACATCGTGACCGTTCGTTG	Intron 2
LpSH1F10	CGTTGTGGGCACTGCACTAGCTTG	
LpSH1R7	GCATACGGTATTTGGAAGAAGAGCTATAGTCAG	
LpSH1R8	GTCACTTTTGGTTGAGTACATCATTGGCATAC	
LpSH1F11	GCTCTTCTTCCAAATACCGTATGCCAATG	Intron 3
LpSH1F12	CCAAATACCGTATGCCAATGATGTACTCAAC	
LpSH1R9	GAAGGAACCCGCTGCCTCTTCTC	
LpSH1R10	GGTTATACGCTGAAGGAACCCGCTG	
LpSH1F13	GAGAAGAGGCAGCGGGTTCCTTC	Intron 4
LpSH1F14	GCAGCGGGTTCCTTCAGCGTATAAC	
LpSH1R11	GCTTCTCTGTGGCTTATGTCAGGGTTG	
LpSH1R12	GCTGCAGTGCTGAAGGCTTCTCTG	
LpSH1F15	CAAACAACCCTGACATRAGCCACAGAG	Intron 5
LpSH1F16	CCTTCAGCACTGCAGCAAAGAACTG	
LpSH1R13	GCTCGCAAGTGATCTATGTACCCACTC	
LpSH1R14	AAGACTTTCATATGCATGGGCGTTAGCTAG	
LpSH1R15	AGGAATTTAATTTTGGCATTATTCAAACCAGTTCCT	
LpSH1I1F1	CGTTGTGTAGAGTGTTGCTATTCTTTATCTGAATC	Middle part of intron 1
LpSH1I1F2	TCCAGGAAACCGATAGGCTTAAACTATTACAC	
LpSH1I1R1	TGTGATCTTCGATGTATAGAGAGAGAGATCGAG	
LpSH1I1R2	ATGCAGGTGTGTTTGTCACCAAGCTAG	

Appendix 3.2 The concentration and purity of total RNA in three biological replicates

	Collection stages	Concentration (ng/µl)	A260/280	A260/230
1st Bio-rep	1-2cm	1060.87	2.14	2.03
	4-8cm	1398.51	2.12	2.14
	>10 cm	1592.81	2.13	2.19
	0 daa	624.11	2.12	1.96
	2 daa	718.75	2.14	1.93
	5 daa	347.94	2.15	1.86
	10 daa	638.87	2.14	1.92
	15 daa	247.06	2.14	1.51
	20 daa	135.63	2.21	1.34
	node	1092.93	2.16	1.67
	seedling root	731.33	2.16	1.45
	Seedling leaves	2890.91	2.05	2.38
	Flag leaves	1821.93	2.11	2.23
2nd Bio-rep	-1 daa	233.3	2.2	1.43
	1 daa	693.69	2.17	1.8
	4 daa	572.27	2.19	1.76
	7 daa	280.85	2.18	1.7
	10 daa	269.11	2.19	1.69
	14 daa	401	2.15	1.88
	18 daa	1115.15	2.17	1.95
	21 daa	722.77	2.16	1.8
3rd bio-rep	1-2 cm	3164	1.987	2.212
	4-8 cm	1817	1.985	2.179
	>10cm	1570	2	2.121
	-2 daa	745	1.99	1.82
	0 daa	383	2	1.5
	4 daa	918	2	1.8
	6 daa	315.6	2.018	1.377
	10 daa	378.4	2.009	1.631
	14 daa	472	2.1	2
	16 daa	490	2.031	1.763
	18 daa	227	2.054	1.333
	seedling leaves	1897	1.97	2.3
	flag leaves	1735	1.9	2.2
	node	454	2.034	1.796
	seedling root	107.6	2.118	0.928

Appendix 3.3 Full sequence of *LpSH1*

5'UTR: 229bp

CCACATGCACGCACACTTGCACTAGCAATTTGCTCCTATGTGTCTCTCTACTCC
ATAATTTCTGTATTACACGTTGTACTGATATAAATATCGGACGGCACTCTGACC
CGCTAGCTAGTGGTCCTTGGTCTCTTCTTCTACTCCGCCATCTTCCTCATCCTCT
CTCCCCTGCCGGCCTTCCCTTTCTTCTTTTGTTCTCGATATCTTGCAACGTCGGG
AGGGGAAAAGA

Exon1: 72nr/24aa

ATGTCGGCACAGATCGCGCCGGCGGAGCATGTGTGCTACGTGCACTGCAACTT
CTGCAACACAATTCTCGCG

Intorn1: 2596bp

GTAATATGCCCCATCTCTCTCTCCTCTCTCTCTCTCTCTCATGATTCATGAGGAG
ATCTACAGAGAGAGATTAGTGGAGCTAATTCTGTAGGATTTGGCTACTTTCTTG
CCGAATTTGTGTTTATCAGAGTTACTGGAAAAACGGCTAAGAGATTTATGAAT
GATTTTGATCCTCTTTTGTTTCGCTACGCCGTGCTCCTTTTGATTTGGCCGCAGC
TGAGTTTTCTTATCCGTTTCTACTTGTTTCAAACGAGACTTCGAGGAAGAAAAA
GAAAGTGAATGTCAGAACATGTCGTGGAAAATCGATCGAGCTGCTTCTCATCG
AATCCATGAAAACCTAGCTATACCTGCTAATCGATTAGTCTCCTGAGCTAATTC
ACCTCTAGGTATGCTCATGTCATTTGTGGCCTTTGATTTGTCTGACTAATTTCTG
AAATTGCATCCCCGCGCCCCCTTTTCTAATTCTTACTGGCCGGGTTTCTTGATGG
AACTTGTTGTTAGATCATCGTTGAAATTGTGAAATTGTGAAGTTCATGCCTAGC
GATCATGAGTCCATGACAAACTGTACTGTGATGTGTGCAGACTTCATGTAGTA
GTACTTGTTATTTTTATTTTCCAAGAACAGCAGCCCTAACCTATTTCAATTGTGG
TGGTTTGATTCTAACTAACTCTATTTGAAGACAATTATGGATTTTTTTCTACTTG
AATATGCAACACAGAGCATGCAATAACAGTCATAACTTCGGTGTATGATCTAT
AAAGTTAGCAAGTTAAGTTATTGATCTAAATAGCAATGAAACGTTGTGTAGAG
TGTTGCTATTCTTTATCTGAATCAAGTCATGTTTTTTTTCCAGGAAACCGATAGG
CTTAAACTATTACACGCAATAGCTAGAACGTAAACTCTTTTATTTATTTATTTAT
TACTTGAACAGAGGAACAGTAGCATACTGTGTTAGCGAGTTGTCACCTAGAAA
TTCTGAGTAATGGATGCTTCCTGTCTTCCGCTCTCTCTCTCTCTATATATATATA
TATCTCCAGTTTGCAGTATGCGTTAATCTGGTTAAAATTGCTACTGGCAATGTG
GAGCTTGTGCCTTGGTCTTTTTCCCTCTTTTTTGGGATTTTTGTTGTCCTTGGCAT
CGAATTAGTACTGTTTGTACTCTTCCTCTCAGTTTCTGCACCTTTATTGTTTGGT
TCGTTCATAAGTACATATATATGCATGACACAAGTACTTGTGTGCTGGAGTAAT
AATTATAGAAAGAAAAAAAGGGAAAAACTAGCTTGATGTACACTATGAACTTG
GAGATTGAAGTAATGCCGTCTTGGTCTCTGCTAACCAAGATAAAAAGGAGGTT
CAGCTGATATATCTCCCACTCACCAAATGGAATGTTAAACTGATTATGCACAGT
GCATCCTGAATTAATTAAACGAGATCTTGCCTTCTCTGGAACTGCAATGCCGTA
TTTGATCCTCTCTCTCTCTCTCTCTCGATCTCTCTCTCTCTACATCGAAGATCAA
AAGGAGAGATGGATCTAGCCTTGTTGACAAACACACCTGCATCTGCGTGACCC
ACTAAGATGTACTATGCACAGTCTAGGAAACGGCAATGCCGGAATTGAGTCTT
TTCTTTTTTTATTCTCGATCTCTCTCTCTATACATCGAAGATCACAAGAGAAGAT
GAATCTAGCTTGGTGACAAACACACCTGCATCTGCGTGACCCACTAAGATGTA
CTATGCACAGTAGTGCATGGTGTGGTGCTTATCGTGAGCTAGTCTAGCTAGCGC
GCGTGAAGTGCTTTTCTGCATGAACCATCGCTGTCATCTTCCTCTCTTTCCTACA
GTGCTTATAGGTGAGATCTTTAATCCTTTTTGCACTGAAGGGAGCATGGAGCAG
ACATGACGGTAGCGAGCATGGCAGGCCATGGCTAATATTCTTTTGAGAAGAAC
CCTCACCATACATAATACTGCACGCATCATTCTCTAAGAAATATTTTTCATATA
TTTTATCTGAAAATAAGCAATTCGAGATATGGTCTGTGATATGAGTCAAGTCCC
TTTCATGGGAGCAAACACAATCATCTGCAGATGCATTTGGCTCACTTGCTTATA

CTACTCTTTCTTATGGGAGTGAGTGCTCAAAAGAAGGAGGTTGAGGAATCAAA
CAAAGTATAGCTAGAGATACGACCAGCTGGTCCTTCTTCAGGTTGTGGTTGCA
GAGCACCCCTGAAAAGAAAACGGCAAGAGTGCAGCCAATGAGCAGGCATCGG
AAAAGGTTGCCCCAGTTCTGATTTGGTCCAGCAGAGCCCTGATAAAAGTCATG
GAGAGCAACTGTTTTACAGCAGGCACGAGTTATTTATATAGAGAGAGAGAGAG
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGCATATATGTGTG
TAGATCTGCACCAGCCTTTTGAAAATGCTTGGTTGGAGTTTAGCTCTTTGGTAC
TTCACTCAGAGAGACCTGGTTATTTTTCTTCTTCTGTTTCTACCTTTTCCCTTCGT
GTTTTCTTCTTCCTTGGATGAACTAAAAAACCTAATTGTGTCTCCTTGTGATCGA
TCTGCAG

Exon2: 126bp/42aa

GTCAGTGTTCCCAGTAATAGCATGCTGAACATCGTGACCGTTCGTTGTGGGCAC
TGCACTAGCTTGCTGTCGGTGAACTTGAGAGGTTTAATCCAGTCACCACCACCT
GTGCAAGATCATTCCCAG

Intron2: 722bp

GTTAACATGCATCGTAGTAATAGTTCATGAAAGATCATGTTTGCATGAAATTAT
CTCACCTAAAGGAAGGTTGAGATGTGTATATATAGCTGTATATATGATATTATA
TGGTGTTATTCCTAGCTTCCTAAGACAGCTGTAGTGGTTTGCTAATTTGATTTA
ATCACATTTCAAAACATTGTGTTATATTTATATTATTGGAGATAATACAGTCAT
GCCTACTGTTTGCATTATTTCTCTCTACTACTACCATAGAGGCTAGAAAGATGG
TAAGTTTCACAACTGGAAGGGAGAGGTGCCAAACCATACACCTGGCTAGAGCC
AAGCTTACCCTAACTCCTACTCTAGCTTATAGCAGAGAAAAGCCAGCATGCAT
CGGTCATGTCATGCTTCAACGCTTTATACAGTAAAGACAAACGGCAGCATGCA
CACTGTTCGCTCACAGGAAAGAGTTTTAATGCTCACATGCAAATACAAAAGTA
GTACTACATCTCTAGTGTCGTGTACCTGATCGTATATTGTTATATTTTTTTGTGC
GGGTATTTTGTTATTTTCTTTTATTGATTTTGAATAGCAATCTTGCACAGCTAGC
TTGCGATGTTTATACATTCCAAATCCTCCCGTGAATATATGTTAGAAGTATTCT
TCTGTTCTATGTTGATAATTAAATGATGCTTCAAGTCAATTAATGAATTAACTT
TATAGTTTTGCTTGATGGCAG

Exon3: 129bp/43aa

GAGAATCTCAAGGCCCACAATATCAGCTTTCGGGGAAATTACCCTGACTATAG
CTCTTCTTCCAAATACCGTATGCCAATGATGTACTCAACCAAAAGTGACCCAGA
ACATATGCTACACATGCGACCA

Intron3: 1792bp

GTAAGATCATGCATATATATTTTGATATTTCACTTTAGTATGTGAATCGTAACT
CTTGTTCTTCCCTCTTGATCAATTTGTATACATAGGTGAATTGCAGTACATGTGT
ATAGTCTTTTCCTTAGGAAAAATACAAAGTTACACCTTGAACTTTTGTACTCTC
AAGCCAACTCATCAAACCGATACAAAAAACCATTTTGGTGGTTTTGAAAAGTG
TTTTGCCACAAATATATAATGTTTTTTTTTCGCGAAAACGCAAAAACCTTGCGT
TTCGATGCATTGATAGAAAAAGAAGGTTATATGTACAAGTCTAAGGGCGGATG
AACACCACACCATACAACTCGACCCAAGAAACAGAAATGCTAATCTAGGGGA
GCAAAAAGCGCATGCGCCCTGGGCACCCGTGTCCGCCCATTGGCGCGCCTCAA
CAAATATATAATGGGTTTCGCCAAACATGTGTCAGAAATTTATCCTGCGAAAA
GTATGTATCCAAAAGTGAAGATTGCCACCTCAAGTTAGGCCAAACGGGGGAGA
TATCGAGCAGAGATGGTACTATTTTGCTGAAATTCTACATGCCAACTATAAATC
AGCCACGTATTCGCCCGCAATATAAAAATGGGTTTGTTGTCTAGTTTTTCAAGT
TCAAGATGTACTATTTGTAGGTTAGGAGTCCAACCATCAATATAATTTGCACCA
TTATCTGAGATTTTTTACAAGAAAATCTTTAAAATTGAGCACGGAAGAATAAA
CTTAATGCAAAGTACGATTGGTGTAGGTATTTACTATAGTAGTTTTTTTTTGA
AAGGTTACTATAGTAGTTTTTATTATGTATACATCTCACATTGCTTTGGTATCTA

ATGAAAATAAAATGGACTCAAGCATAGAGAAAAATGAATAGAGAAATGAAAC
TTAAATTTCCTATAAAATATTTTGTAGTACACTTTCATTAACTATTAAGTCCCAT
TAAGGACATGATCTGAATATCTTGATTCGTGCAAGTCCAAATGGATACAAGAA
AAAAAAAGTTACAAATTGATGAAAATATCTTAATAAATTGACTACCGTTATCC
ACTTTTAAATCACTGTGAGATATGGTCACCGACTTAGCTATTTATGTGTCGGGC
ATACTCCTTTGAGAAAATACATACCTCTTGCATGAATACGTAAACTTGTCTTTC
ATGGGTAAAAACATTATACCGCTAGGCATAATCATTATTAAATAACTAACTTTG
AGTTTCCATAATATCTAGTGATAATTATTAGGGATTAGTGTGTTTCTAGATATC
TATTGTGTTACATCTTGATTAATTTTAGCAAAAGTATGTATGTCCTCCATCTTGG
GTGCCGAAGTCTAGAAATGATCCTCAATTCTAAAACCAAGAAAACCAGCTCAC
TCGATTTGTAAAAACGAATTAATTTGGTTCCTTGTATTTTTTTGGAGTGGTTTTG
GTGATCATGTGTCCTTGTTTTGACCAAAACCTAATATAATTAATGGCAAATTTT
GACTGTTGACCTTGCCATATGTGCATGCATGTCAACACTTGGCTGATCGAAATC
AAAACTAGGCCACTCATGCCAAAACCATAGAAATTTGCATGACGGATTAAGAG
TATGCAGTTTTAGAAGCGTAAGGACATACAGTTTTTTTATTTAGGAGCATAGTA
TATTTCTACACTTACCAATTATAGTTTTTTCAAAGGGAGATACATTTATATCAA
GATTCCAACAGTTTGGATAAAAATATATGTGTTTCCCTTAAATTTATGAGGATG
AATCTTTGATACTGCAGCA

Exon4: 48bp/16aa

GCTACCGAGAAGAGGCAGCGGGTTCCTTCAGCGTATAACCGATTTATT

Intron4: 98bp

AAGTAAGTTTCAATCAATTAAATGCAAGCTTGTACTAAACATAATGAAGCGG**T**
ATTGTCCTCTACTAACCATATATACCATTCTGACATATAAATTTC

Exon5: 78bp/26aa

AAGGAAGAGATACGGAGGATAAAAACAAACAACCCTGACATAAGCCACAGAG
AAGCCTTCAGCACTGCAGCAAAGAAC

Intron5: 436bp

GTTAGTACATGCTACGCATAACACCTATGTTGTACTACCACAGTTTTACATTTG
CAGGTCACAACTACTCCAGTTTCGTCCGTCAAAAAAACAAAAAAAAAACTACT
CCAGTTTCCACACTTGACCATAACCCATAGTTTTGGTATATAACAAACAACTTT
CATATGCATGCATATGCAAAAGAGCCACAAATTAATAGCTAAATCACTCAGAT
AGAAATTGACGCGTATGTTGCTCTGTCATTTGCCCTGGCCAAAGCTTATATTAC
TGCAGATAACTCCATTTCTAGGGCATGTTTGGAATACAAACAGAAAATTCCTGT
AGATCTCAGATTTTTTTTGTACCTTGTCCCAAAGGGATTCTAGTGTTTCTCTGGA
GACAGTTGATACGAAATACTACTGTTTAGTTTACATCTCTCTACTGATTTCTATT
TCAG

Exon6:111bp/37aa

TGGGCGCATTTCCCTAACATACATTTCGGCCTAGGCTCCAATGAGAGCAGCAA
GAAGCTGGACGAGGCCATCGCGGCGCCTATCCCCCAGAAAGTTCAAGGTCTCT
ACTGA

3'UTR: 200bp

GACAAATATTCCACTATGCATGATATTTCATATTGTAAGTAATTTAAGGAAATC
TAAGTTATTTCTACATCAGTTTGTGAGTGCGTACATAGATCACTTGCGAGCGGC
ATGCGTGAAGTGAGAGAAAAAAGCATATGGATGGCTTCTTAATAAATCCCTTT
GATCGATCATCCTGTTCTTATCCTGTTTTTTTCTATTAGGAACTGGTTTGAATAA
TGCCAAAATTAAATTCCAATCTCTAGAGGATCCCCGGGTACCGAGCTCGAATC
GTAATCATTCAATGTTCCC

Appendix 3.4 A brief outline of the transgenic study

To investigate the function of *LpSH1* in perennial ryegrass, a transgenic study was carried out to up-regulate and down-regulate the expression level of *LpSH1* in rice (*O. sativa* Nipponbare) and perennial ryegrass (cv. Nui). Three expression vectors were constructed, including two vectors for overexpression and one for RNAi.

- **Vector construction**

To generate the DNA constructs, p35S::LpSH1-GFP, full length coding sequence (CDS) of *LpSH1* was amplified with the primers shown in Table. The PCR products were then cloned into the binary vector, pSCZ3, which has a GFP insertion. In case GFP gene affects *LpSH1* gene function, another overexpression vector without GFP, p35S::LpSH1, was constructed using pSCZ3 as well. To generate the *LpSH1* RNAi construct, a 200 bp of *LpSH1*, spanning exons 3 to 6, was amplified with the primers shown in Table below, and the PCR product were inserted into the LH-FAD2-1390 RNAi vector.

- **Plant transformation**

The two *LpSH1*-overexpression vectors were transformed into non-shattering rice (*O. sativa* Nipponbare) by *Agrobacterium*-mediated transformation. Due to contamination of the transformed calli, no plant regeneration has been obtained. For the RNAi, the RNAi vector plasmid was transferred into multi-shooting callus induced from perennial ryegrass cv. Nui seeds. The rice and ryegrass transgenic studies failed, because the transformed callus did not regenerate.

Subsequently, two vectors, *LpSH1* RNAi and p35S::LpSH1-GFP, were sent to a biotechnology company (Doublehelix Biology Science and Technology Co., Wuhan, China) to perform custom ryegrass transformation. The transformed callus from ryegrass seed still cannot be regenerated.

Recently, the vector, p35S::LpSH1-GFP was sent to Doublehelix Biology Science and Technology Co., to be transformed into rice callus by *Agrobacterium*-mediated transformation. This process is on-going.

Primers used for construction of transgenic vectors

LpSH1 overexpression	LpSH1-OE-GFP-F (Afl II)	cgg**CTTAAG**ATGTCGGCACAGATCGCGCCGGCGGAGC
	LpSH1-OE-GFP-R (BamH I)	cg**GGATCC**GTAGAGACCTTGAACTTTCTGGGGGATAG
	LpSH1-OE-F (Afl II)	cgg**CTTAAG**ATGTCGGCACAGATCGCGCCGGCGGAGC
	LpSH1-OE-R (Xba I)	gc**TCTAGA**TCAGTAGAGACCTTGAACTTTCTGGGGGA
LpSH1 RNAi	LpSH1-RNAi-F1(*KpnI*):	gg**GGTACC**AGAACATATGCTACACATGCGACCAG
	LpSH1-RNAi-R1(*SacI*):	c**GAGCTC**GAGGAATATTCGTCTCAGTAGAGACCTTG
	LpSH1-RNAi-F1'(*MluI*):	cg**ACGCGT**AGAACATATGCTACACATGCGACCAG
	LpSH1-RNAi-R1'(*BamHI*):	cg**GGATCC**GAGGAATATTCGTCTCAGTAGAGACCTTG

Appendix 4.1: Genotyping in perennial ryegrass population

4.1.1 Materials and methods

4.1.1.1 Initial selection of LpSSR primers using agarose gel electrophoresis

Twenty-six pairs of LpSSR marker primers were selected from previous studies (Jones et al., 2001; Kubik et al., 2001; Wang et al., 2009; Liu et al., 2010) and synthesised by Macrogen Inc. South Korea. All LpSSR primers sequences are listed in Table 1.

Table 1 LpSSR primer sequences

Name	Forward primer	Reverse primer	Original name	Expected amplicon sizes (bp)
LpSSR1	TGACTTCTCTCGATCCT	ATGTGACTACAAAACCA	LP20	*
LpSSR2	CGCAGCTTAATTTAGTC	GCTTTGAGTATGTAAAGTT	LPSSRH01A10	103-109
LpSSR3	CTAATCTGGGTGCGTTTGC	CGTGGACGAAGGTAGGCTTG	LPSSRh01c07	245
LpSSR4	ATTGACTGGCTTCCGTGTT	CGCGATTGCAGATTCTTG	LPSSRH01H06	150
LpSSR5	CGGCCACCCTTGATAGAG	TCGTCAAGGATCCGGAGA	LPSSRH02C11	198
LpSSR6	TGCCACAGTTGTTCTCCTCTTGAGG	CAAATATGGCTCCTACGCTGCGCTT	LPSSRh10g02	96
LpSSR7	CTAAGGGAAGACATGCACTGAGGC	GGCAGAGGAAAGAAGGGCAAC	LPSSRh12g06	371
LpSSR8	GCCCAACCTCGAAGACTACC	TCACATCGACGATGGCACCC	LPSSRk02d08	310
LpSSR9	TCTGAAAGCCCGAGTGAGCG	CGACTGTGGCAGGGATGACG	LPSSRk02e08	174
LpSSR10	GGGAATCTGGCAGAAGTATCACGT	GAAGATCTGGCCAAGTCTAATCCG	LPSSRk03b03	296
LpSSR11	ATGTGCCAATGGAAAACTTGAGGA	TGCCAAAGTTAGGCTTGGCAAAGT	LPSSRk05h02	202
LpSSR12	AAGGAGACCTGGCAGCTTGGTGCT	CGACAAAGGTTCACAATGGAGGAG	LPSSRk10h05	282
LpSSR13	GGTCTGGTAGACATGCCTAC	TACCAGCACAGGCAGGTTC	M15-185	142-198
LpSSR14	TGCTGTGGCTCTTGTGAC	AGCCGAGGCTCAGCTCGA	M16-B	137-178
LpSSR15	GCAACTTCTATCGAGTTG	GAGGCTCGATCTTCACGGA	M2-148	*
LpSSR16	AGAGACCATCACCAAGCC	TCTGGAAGATTTCCTTG	M4136	183-207
LpSSR17	TCTGCATTCGTTGTCTCACTG	GAGCCGTCGCACCCCTG	PR37	100-124
LpSSR18	CATTCATCCACGTTAGAC	CTTCCACGACTGCTTC	PR39	97-151
LpSSR19	TGAGCACCATGAAGGAG	GGTTGTCCGCAGGTATT	LPSSR017	224-254
LpSSR20	GGGGAAATACAGTTCTGC	GATGCTCCTGCCTACTTTA	LPSSR020	225-297
LpSSR21	CGATGAACTCAAGGGGGATT	GCACCGGTCTAGGGACAGAA	LPSSR058	322-350
LpSSR22	GCCAGTGCCCATTCCGATAA	CCCCACTCCAACCAAAGCAA	LPSSR066	233-302
LpSSR23	GACCCCGAGACAGCCTA	ACGCATATGGTCTTCAGAA	LPSSR112	234-264
LpSSR24	AAAGACCGCATACGAAGT	AACCAAAGCCTCAAGACA	LPSSRH01A02	146-150
LpSSR25	TGGAGGGCTCGTGGAGAAGT	CGGTTCCCACGCCTTGC	LPSSRH01A07	77
LpSSR28	TGACCAGCTTGAAATCCTCATCTT	ACCAGCGCCACCATCACCTC	LPSSRk03g05	185

* amplicon size is not annotated in the original article.

Genomic DNA was extracted from leaf tissue of 140 individual perennial ryegrass plants from seven cultivars (20 of each cultivar, A: RI009, B: 8781B, C: 8719A, D: R3007, E: MTX608, F: Nui, G: R3021) by using the method described in Chapter 2. All extracted genomic DNA was assessed by using the NanoDrop ND-1000 spectrophotometer and then diluted to about 200 ng/µl with Milli-Q water, then stored at -20°C.

The PCR amplifications were performed with LpSSR marker primers, using mixed genomic DNA from four cultivars (A, B, C, D). Each reaction contained 1.5 µl 10x PCR buffer (Bioline, UK), 0.2 mM of each dNTP (Bioline, UK), 1 Unit of Taq DNA polymerase (Bioline, UK), 0.66 µM of each primer and about 200 ng of template DNA in a total volume of 15 µl. The reaction was subjected to 94°C for 5 min, followed by 35 cycles of 94°C for 40

s, 45°C annealing for 30 s, and 72°C extension for 40 s, and final 72°C extension for 6 min. The reaction was also performed at different annealing temperatures 50, 55, 60, 65°C, respectively, as shown in Table 2. The amplified products were separated on 1% (w/v) agarose gel (Invitrogen, USA) with TAE buffer at 6 V/cm for 1 h.

Table 2 LpSSR marker primers initial selection outline

Template DNA	Ta (°C)	Primers
Mixed genomic DNA from Cultivars A, B, C, D [a]	45	All LpSSR primers
	50	
	55	
	60	
	65	

Ta: Annealing Temperature

[a] The mixture of genomic DNA contains all of four cultivars genomic DNA; A: RI009, B: 8781B, C: 8719A, D: R3007

4.1.1.2 Second selection of LpSSR primers among perennial ryegrass cultivars

After the first screening, 18 selected primers were used for cultivar identification. The selected primers were used to distinguish the four cultivars (A, B, C, D). Genomic DNA from ten individuals from one cultivar was mixed as template DNA for the following PCR. The 15 µl of reaction master mix contained 1 Unit of Taq DNA polymerase (Bioline, UK), 1.5 µl Bioline 10x PCR buffer, 0.2 mM of each dNTP (Bioline, UK), 0.66 µM of each primer and about 200 ng of template DNA. The reaction was subjected to 94°C for 5 min, followed by 35 cycles of 94°C for 40 s, 50°C annealing for 30 s, and 72°C extension for 40 s, and final 72°C extension for 6 min. The reaction was also performed at different annealing temperatures 55°C, 60°C, 65°C, respectively (Table 3). The amplified products were separated on 1% (w/v) agarose gel with TAE buffer at 6 V/cm for 30 min.

Table 3 Second selection of LpSSR primers between cultivars (cultivar identification)

Template DNA	Ta (°C)	Primers
Mixed genomic DNA from Cultivar A, B, C, D [a]	50	20
	55	3, 4, 5, 6, 7, 9, 11, 12, 13, 14
	60	4, 5, 6,7 9,10 11, 12,13 14 15 21 22 23
	65	8, 28

Ta: Annealing Temperature

[a] The four mixtures of template DNA contains ten individuals from each cultivar; A: RI009, B: 8781B, C: 8719A, D: R3007

4.1.1.3 Genotyping within cultivars using agarose gel electrophoresis

Genomic DNA from 140 individuals from seven cultivars (20 individuals of each cultivar, A, B, C, D, E, F, G) was subjected to the following PCR with selected primers, including LpSSR 10, 21, 22, 23. Each reaction contained 1 Unit of Taq DNA polymerase (Bioline, UK), 1.5 µl Bioline 10x PCR buffer, 0.2 mM of each dNTP (Bioline, UK), 0.66 µM of each primer and about 200 ng of template DNA in a total volume of 15 µl. The reaction was subjected to 94°C for 5 min, followed by 35 cycles of 94°C for 40 s, 60°C annealing for 30 s, and 72°C extension for 40 s, and a final 72°C extension for 6 min. The amplified products were separated on 3% (w/v) agarose gel with TAE buffer at 6 V/cm for 1 h.

4.1.1.4 Reassessment of LpSSR marker primers for HRM analysis

After initial primer selection by standard PCR and assessment by agarose gel electrophoresis (Section 4.2.1.1), the selected 18 pairs of primers (Table 12) were used for cultivar identification and genotyping within a cultivar. Here, the 18 pairs of LpSSR primers were reassessed to determine if these primers could be applied to HRM analysis. Genomic DNA from 18 individuals from perennial ryegrass cv. Glencar (tetraploid, labelled as W) was extracted, and genomic DNA from two individuals (labelled as W3 and W18) was used as template DNA.

Four HRM master mixes with EvaGreen dye (20x in water, Biotium, USA) were used to assess the utility of the LpSSR primers for HRM analysis, since HRM analysis requires high specification PCR amplification and a reasonable length for resolution. Four HRM master mixes included two homemade master mixes (Roche Taq DNA polymerase and Bioline Taq DNA polymerase) and two commercial HRM kits (Qiagen Type-it HRM kit and Biotium HRM kit), as shown in Table 4. The details of the amplification reactions and cycling programs for each master mix are listed in Table 5-11.

Table 4 LpSSR primers used for the optimisation of homemade mister mixes and commercial kits for HRM analysis

Master mix	Template DNA	Selected primers
Roche Taq DNA polymerase	Genomic DNA from W3 and 18 [a]	3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 21, 22, 23, 28
Bioline Taq DNA polymerase		
Type-it HRM kit		
Biotium HRM kit		

[a] Genomic DNA from two individuals W3 and 18 were applied to PCR separately. W: cv. Glencar.

Table 5 Roche Taq DNA polymerase

Component	Volume per 15-µl reaction (µl)	Final Conc.
Roche 10x PCR buffer	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
25 mM $MgCl_2$	1.8	3 mM
Taq DNA Polymerase	0.12	0.6 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	2	0.66 µM
EvaGreen	0.75	1X
PCR-grade water	Make up to 15 µl	

Table 6 Cycling protocol for HRM analysis with Roche Taq DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Initial denaturation	95°C	2 min
Cycling, 40 cycles	Denaturation	95°C	10 sec
	Annealing	58°C	15 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	70-90°C, 0.1°C increments	2 sec

Table 7 Bioline Taq DNA Polymerase

Component	Volume per 15-µl reaction (µl)	Final Conc.
Bioline 10x PCR buffer	1.5	1x
2 mM dNTP mix	1.5	0.2 mM each
50 mM $MgCl_2$	0.5	1.6 mM
Taq DNA Polymerase	0.2	1 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	2	0.66 µM
EvaGreen	0.75	1X
PCR-grade water	Make up to 15 µl	

Table 8 Cycling protocol for HRM analysis with Bioline Taq DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Initial denaturation	95°C	5min
Cycling, 40 cycles	Denaturation	95°C	10 sec
	Annealing	58°C	15 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	70-95°C, 0.1°C increments	2 sec

Table 9 Type-it and Biotium HRM PCR Master Mix

Component	Volume per 15-µl reaction (µl)	Final Conc.
2X HRM PCR Master Mix	7.5	1x
Template DNA	1	
5 µM Primer Mix	2	0.66 µM
PCR-grade water	Make up to 15 µl	

Table 10 Cycling protocol for HRM analysis with the Type-it HRM kit

Stage	Step	Temp.	Time
Initial Denaturation	Initial denaturation	95°C	5min
Cycling, 40 cycles	Denaturation	95°C	10 sec
	Annealing	58°C	15 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	70-95°C, 0.1°C increments	2 sec

Table 11 Cycling protocol for HRM analysis with the Biotium HRM kit

Stage	Step	Temp.	Time
Initial PCR step	Denaturation/ Activation	95°C	2 min
Cycling, 40 cycles	Denature	95°C	5 sec
	Annealing/Extension	60°C	30 sec
HRM	Fluorescence data acquisition	70-90°C, 0.1°C increments	2 sec

4.1.1.5 Trial of the LpSSR4 and 9 primers for genotyping using HRM analysis

HRM analysis was performed with genomic DNA from nine individuals from cv. Glencar (tetraploid, W) and the LpSSR4 primer. The PCR amplification was performed by using the Type-it HRM kit with three replicates and Roche Taq homemade HRM master mix with two replicates. Reaction components and cycling programs are referred to Tables 5, 6, 9, and 10.

To determine the reliability of the HRM analysis, the perennial ryegrass cv. Nui (diploid), the LpSSR9 primer and a new brand Taq DNA polymerase were used for PCR amplification. Genomic DNA from 20 individuals from cv. Nui was extracted by using the method described in Chapter 2 and the isolated genomic DNA was diluted to about 200 ng/µl as template for HRM analysis. The LpSSR4 primer was used for the PCR using Roche Taq homemade HRM master mix and the Type-it HRM Kit, and the LpSSR9 primer was used for PCR using AmpliTaq Gold 360 DNA Polymerase (Applied Biosystems, USA), Roche Taq DNA polymerase and the Type-it HRM Kit. All HRM results were analysed by using the unsupervised mode of Rotor-Gene ScreenClust HRM Software (Qiagen), which is a tool for analysis of high-resolution melting data from the Rotor-Gene Q.

4.1.1.6 Sequencing of LpSSR4 PCR products

As too many genotypes from LpSSR4 were detected by HRM analysis from cv. Nui and Glencar, it was necessary to sequence the PCR products from LpSSR4 and to confirm the genotyping results from the HRM analysis. Hence, genomic DNA from another 20 individuals from cv. Glencar was extracted. HRM analysis was performed with genomic DNA from 29 individuals from cv. Glencar and the LpSSR4 primer using the Type-it HRM kit. The HRM results were analysed by the unsupervised mode of ScreenClust. The PCR products from 12 individuals, representing different genotypes, were selected for sequencing following cloning by using the TOPO TA Cloning Kit (Invitrogen, USA) following the manual. The constructed plasmids were transformed into *E. Coli* DH5α competent cells. The plasmids were isolated from selected transformed cells by using DNA-spin Plasmid DNA Purification kit (iNtRON Biotechnology Inc. South Korea) following the manual. All plasmids were sent to Macrogen for sequencing. The sequencing results were aligned and analysed.

4.1.2 Results

To reduce the workload in mutation screening and increase the probability of identification of useful mutations, it was important to select different genotypes with different genetic backgrounds from different perennial ryegrass cultivars. A method for genotyping using LpSSR markers was optimised.

4.1.2.1 Initial selection for LpSSR primers by agarose gel electrophoresis

The PCR products from 26 pairs of LpSSR primers and the mixed genomic DNA from four cultivars as template, at different annealing temperatures, were checked on 1% agarose gel with TAE buffer at 6 V/cm for 1 h. The PCR products annealing at 60°C are shown in Figure 1. The range of annealing temperatures for each primer is shown in Table 12. Eighteen Primers were then selected for ryegrass cultivar identification and genotyping within a cultivar, because the PCR products of these primer present bright band at correct size corresponding to original articles.

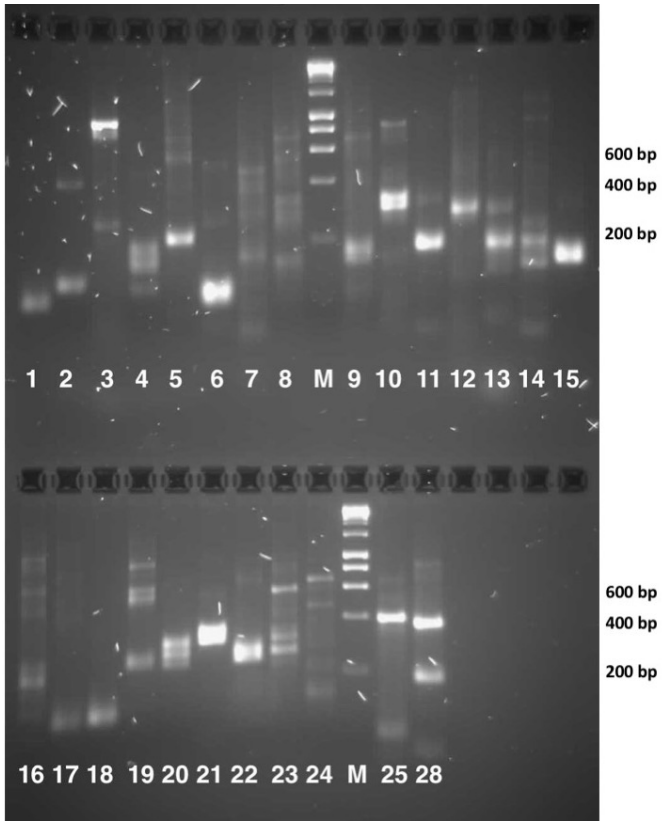


Figure 1 The PCR products of LpSSR1-25, 28, annealing at 60°C from mixed genomic DNA of four cultivars. All PCR products were checked on a 1% (w/v) agarose gel with TAE buffer at 6 V/cm for 1 h. M: HyperLadder I.

Table 12 The 18 selected LpSSR primers and annealing temperature.

Primer	Ta (°C)	Primer	Ta (°C)
LpSRR03	55	LpSRR12	55-65
LpSRR04	60	LpSRR13	60
LpSRR05	55-60	LpSRR14	60
LpSRR06	55-60	LpSRR15	50-60
LpSRR07	60	LpSRR20	50
LpSRR08	65	LpSRR21	50-60
LpSRR09	60-65	LpSRR22	50-60
LpSRR10	55-65	LpSRR23	60
LpSRR11	55-65	LpSRR28	65

Ta: annealing temperature

4.1.2.2 Selection of LpSSR primers for distinguishing among cultivars

The PCR products amplified from genomic DNA from cultivars A, B, C, D. For each cultivar, the genomic DNA was a mix of genotypes and the selected primers were checked on an agarose gel. After second selection among cultivars, the primers LpSSR 10, 20, 21, 22,

and 23 could be used for genotyping, because these PCR products from different cultivars can be distinguished by banding patterns visually. For example, there is a significantly different binding pattern of the PCR products from LpSSR23 (Figure 2), so LpSSR23 could be used for cultivar identification. There are slight differences in product sizes from LpSSR21 and 22 (Figure 2), which means that they could be used to identify genotypes within a cultivar, even though they could not differentiate among cultivars.

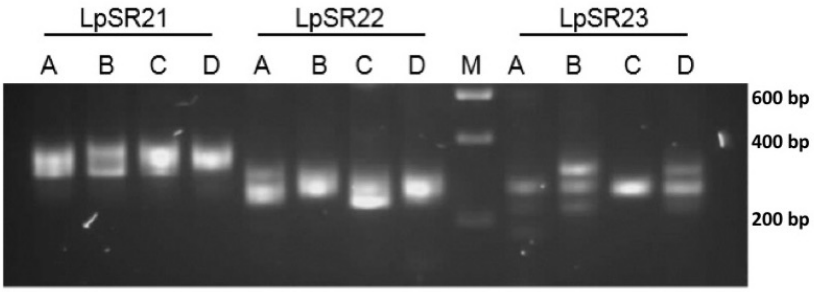


Figure 2 The PCR products of primers LpSSR21, 22, and 23 among the four cultivars, A, B, C, D. All PCR products were checked on a 1% (w/v) agarose gel with TAE buffer at 6 V/cm for 30 min. M: HyperLadder I.

4.1.2.3 Genotyping within cultivar using agarose gel electrophoresis

Based on previous primer selection, 140 individuals from seven cultivars (Section 4.1.1.1) were genotyped using LpSSR10, 21, 22, 23 primer pairs. Figure 3 shows that PCR products of primer LpSSR22 and genomic DNA from 10 individuals from cultivar D (D1-10) are different. About five banding patterns can be observed in Figure 3, and each banding pattern represents one genotype, so at least five genotypes can be identified by using LpSSR22 within the 10 individuals from cultivar D. Gel photos from LpSSR10, 21, 22, 23 PCR products are not shown.

However, it was hard to differentiate the genotypes within a cultivar by visualising band sizes on agarose gels, as the resolution of agarose gel is limited. Consequently, HRM analysis was optimised for genotyping of perennial ryegrass cultivars.

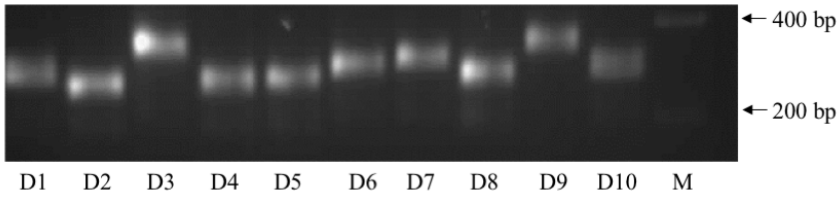


Figure 3 LpSSR22 PCR products from genomic DNA from ten individuals (D1-10). All PCR products were checked on a 3% (w/v) agarose gel with TAE buffer at 6 V/cm for 1 h. M: HyperLadder I.

4.1.2.4 Assessment of LpSSR primers by HRM analysis

The LpSSR marker primers were assessed by using HRM analysis with two homemade master mixes and two commercial HRM master mix kits. Based on the melting curve analysis (Figure 4), the LpSSR primers with the sharpest peaks were selected. Comparing the Roche Taq homemade master mix and Bioline Taq homemade master mix, the product peaks from Roche Taq homemade master mix were sharper than the peaks from the Bioline Taq homemade master mix with the same PCR conditions, such as primers and template DNA (Figure 4-A, B). The shape of the melting curve reflects the PCR product specificity, as the melting curves are generated by the denaturing temperature of PCR product, so the Roche Taq homemade master mix is better than Bioline Taq homemade master mix for genotyping with the selected LpSSR markers. For melting curves from two commercial HRM kits, Type-it HRM kit and Biotium HRM kit, the Type-it kit showed a better result than the Biotium kit (Figure 4-C, D), because of its shaper peak patterns. The different numbers of selected primers for each master mix, including two homemade and commercial kits, are listed in Table 13. The rest of primer were not selected for subsequent HRM analyses as their melting curves did not have sharp peaks, which means that these primers were not ideal for HRM analysis (Data not shown).

Table 13 The selected ideal LpSSR primers for different master mixes including homemade and commercial kits for HRM analyses.

Master mix	Selected ideal LpSSR primers
Roche Taq homemade master mix	4,9,11,17,19,20,23
Bioline Taq homemade master mix	4,6,9,12,17,28
Type-it HRM kit	4,5,6,9,11,12,14,17,28
Biotium HRM kit	4,6,9,11,12,20

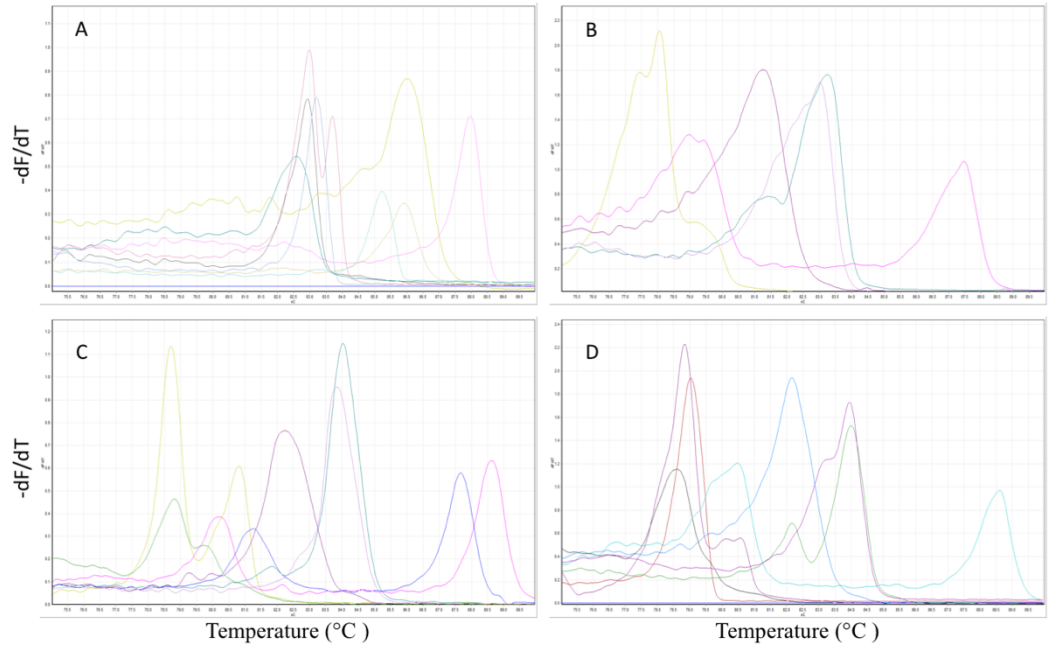


Figure 4 Melting curves of selected LpSSR primers from four master mixes. A: Roche homemade master mix; B: Bioline homemade master mix; C: Type-it HRM Kit; D: Biotium master mix kit.

4.1.2.5 Genotyping with selected LpSSR primers within a cultivar

The HRM results using LpSSR4 primer for cv. Glencar, performed by Roche Taq homemade master mix and the Type-it HRM kit, were analysed by using ScreenClust software in unsupervised mode. HRM melt curves can have different starting points and the scale of each melt curve can be different. Comparison can only occur if all samples are on the same scale, so each curve needs to be normalized. Clusters analyzed in the ScreenClust HRM Software are groups of samples with the same melt characteristics, so each cluster may represent one genotype.

After analysis using the ScreenClust HRM Software, five genotypes were detected by the Roche Taq homemade master mix but six genotypes were detected by the Type-it HRM kit (Table 14). However, three samples, highlighted in Table 14, are classified into different clusters in the two replicates of Roche Taq homemade master mix, whereas with the Type-it HRM Kit, the three replicates show the same cluster results. In addition, in the normalised fluorescence graphs from both master mixes (Figure 5), it is observed clearly that the repeatability of the Type-it HRM Kit is better than the Roche Taq homemade master mix.

Therefore, the Type-it HRM kit is more reliable, and the Roche Taq homemade master mix still needs to be improved.

Table 14 The genotyping results of LpSSR4 within cv. Glencar (W) by using Roche Taq homemade master mix and the Type-it HRM Kit. Different cluster results using Roche Taq homemade master mix are highlighted in yellow.

Template DNA	Roche homemade master mix		Type-it HRM kit		
	Cluster		Cluster		
	Replicate 1	Replicate 2	Replicate 1	Replicate2	Replicate3
W-1	Cluster 5	Cluster 1	Cluster 4	Cluster 4	Cluster 4
W-2	Cluster 5	Cluster 5	Cluster 2	Cluster 2	Cluster 2
W-4	Cluster 3	Cluster 3	Cluster 1	Cluster 1	Cluster 1
W-5	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
W-6	Cluster 3	Cluster 3	Cluster 1	Cluster 1	Cluster 1
W-7	Cluster 5	Cluster 4	Cluster 3	Cluster 3	Cluster 3
W-8	Cluster 3	Cluster 3	Cluster 5	Cluster 5	Cluster 5
W-9	Cluster 4	Cluster 5	Cluster 2	Cluster 2	Cluster 2
W-10	Cluster 2	Cluster 2	Cluster 6	Cluster 6	Cluster 6

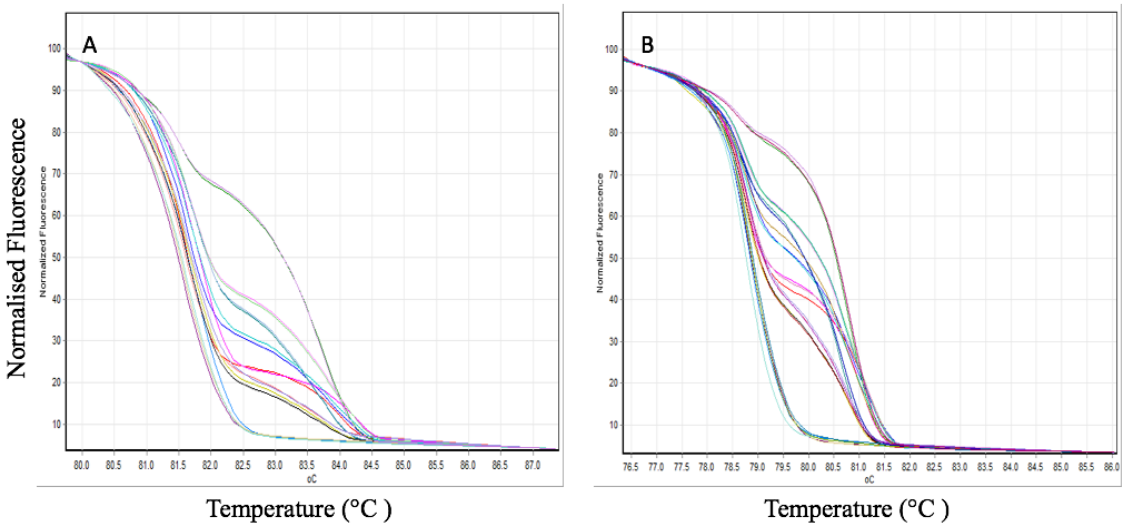


Figure 5 Normalised fluorescence of HRM analysis from LpSSR4 with genomic DNA from nine individuals from cv. Glencar. A: Roche Taq in homemade master mix with two replicates; B: The Type-it HRM kit with three replicates. Each line in the graph represents one sample, and the reproducibility is reflected by overlapping curves.

To further assess the reliability of the HRM analysis, genomic DNA from 20 individuals from cv. Nui was used with the two primer pairs, LpSSR4 and LpSSR9, using AmpliTaq homemade master mix, Roche Taq homemade master mix and the Type-it HRM kit, respectively. For genotyping results from LpSSR4 (Table 15 and Figure 6), the Amplitaq homemade master mix and the Type-it HRM kit present very similar cluster results showing

eight genotypes. For the genotyping results from LpSSR9, the Type-it HRM kit and AmpliTaq homemade master mix present the same cluster results, but the cluster result from Roche Taq homemade master mix show some differences, highlighted in Table 15. In the normalised fluorescence graphs of LpSSR9 (Figure 7), the curve patterns from Roche Taq homemade master mix are also different, compared to the other two master mixes. Based on these results, the AmpliTaq master mix performed better than the Roche Taq homemade master mix. As the genotypes distinguished using LpSSR4 are highly variable in both cultivars, Nui and Glencar, so the sequencing of PCR products from LpSSR4 was carried out.

HRM analysis was developed as an efficient and cost-effective method to detect DNA sequence variations including SNPs in plants and animals and this method has been applied to genotyping by SSR markers in several plant species (Mackay et al., 2008; Mader et al., 2008; Muleo et al., 2009; Wu et al., 2010; Arthofer et al., 2011). In this section, HRM analysis was examined as an alternative to the gel-based platform for genotyping using SSR markers from perennial ryegrass. The primer pairs, LpSSR4 and LpSSR9, were employed for genotyping within two different cultivars: diploid cultivar Nui and tetraploid cultivar Glencar. After analysis by ScreenClust, eight clusters were detected by LpSSR4 in cultivar Nui, but only three clusters when LpSSR9 was used (Figures 4.5-4.7). This result is in agreement with Kubik et al. (2001). They demonstrated that each SSR locus in each cultivar has different degrees of heterozygosity and different numbers of alleles. As a consequence, a reliable genotyping result is decided by the number of loci used and the number of individuals sampled per cultivar (Kubik et al., 2001). However, the two LpSSR markers presented very different genotyping results that the number of genotypes from LpSSR4 was almost three times greater than LpSSR9. This great variety increased the difficulty of genotyping. Also, the same primer pair, LpSSR4 and LpSSR9, but with different HRM master mixes, gave different genotyping results (Figures 4.5-4.7), indicating that the choice of HRM master mix is a foundation for reliable results.

Table 15 The genotyping results of LpSSR4 and LpSSR9 within cv. Nui by using the Type-it kit, Amplitaq homemade master mix and Roche Taq homemade master mix (only for LpSSR9). Different cluster results from the same primer pair but different master mixes are highlighted in yellow.

Template DNA	LpSSR4		LpSSR9		
	Type-it HRM kit	Amplitaq homemade master mix	Type-it HRM kit	AmpliTaq homemade master mix	Roche homemade master mix
Nui 1	Cluster 3	Cluster 2	Cluster 3	Cluster 3	Cluster 3
Nui 2	Cluster 8	Cluster 8	Cluster 3	Cluster 3	Cluster 2
Nui 3	Cluster 4	Cluster 6	Cluster 3	Cluster 3	Cluster 2
Nui 4	Cluster 8	Cluster 8	Cluster 3	Cluster 3	Cluster 2
Nui 5	Cluster 7	Cluster 5	Cluster 3	Cluster 3	Cluster 2
Nui 6	Cluster 4	Cluster 6	Cluster 1	Cluster 1	Cluster 1
Nui 7	Cluster 6	Cluster 3	Cluster 1	Cluster 1	Cluster 1
Nui 8	Cluster 8	Cluster 8	Cluster 3	Cluster 3	Cluster 3
Nui 9	Cluster 8	Cluster 8	Cluster 1	Cluster 1	Cluster 1
Nui 10	Cluster 8	Cluster 8	Cluster 1	Cluster 1	Cluster 2
Nui 11	Cluster 2	Cluster 7	Cluster 1	Cluster 1	Cluster 1
Nui 12	Cluster 7	Cluster 5	Cluster 3	Cluster 3	Cluster 3
Nui 13	Cluster 7	Cluster 5	Cluster 3	Cluster 3	Cluster 3
Nui 14	Cluster 8	Cluster 8	Cluster 1	Cluster 1	Cluster 1
Nui 15	Cluster 8	Cluster 8	Cluster 1	Cluster 1	Cluster 1
Nui 16	Cluster 5	Cluster 1	Cluster 3	Cluster 3	Cluster 3
Nui 17	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 3
Nui 18	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 3
Nui 19	Cluster 5	Cluster 1	Cluster 2	Cluster 2	Cluster 2
Nui 20	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 3

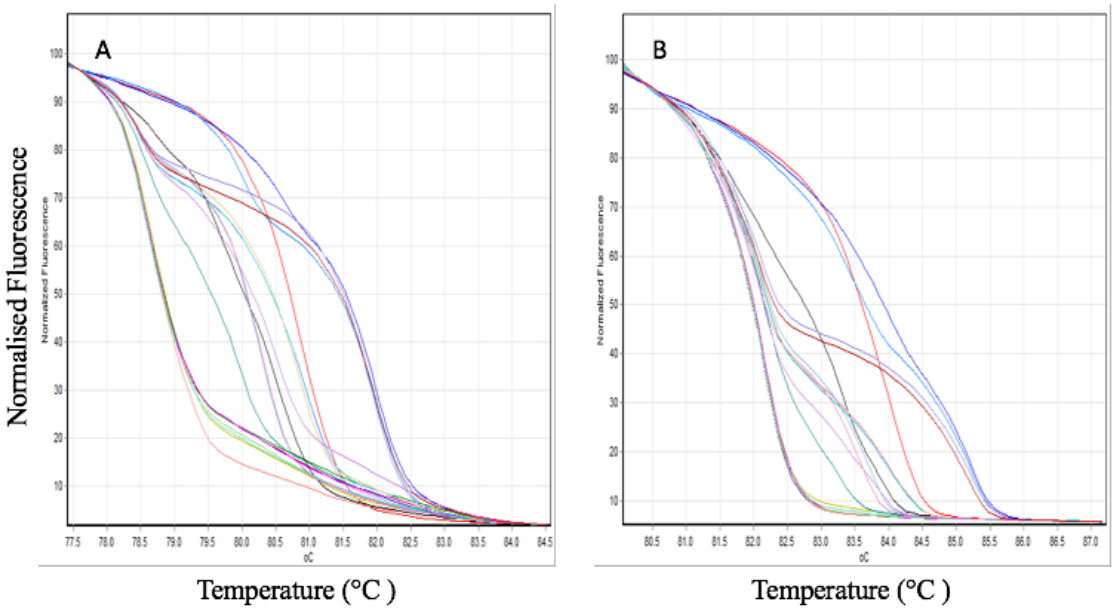


Figure 6 Normalised fluorescence graphs of HRM analysis from LpSSR4 from genomic DNA from 20 individuals from cv. Nui. A: Type-it HRM Kit; B: AmpliTaq homemade master mix. Each line in the graph represents one sample.

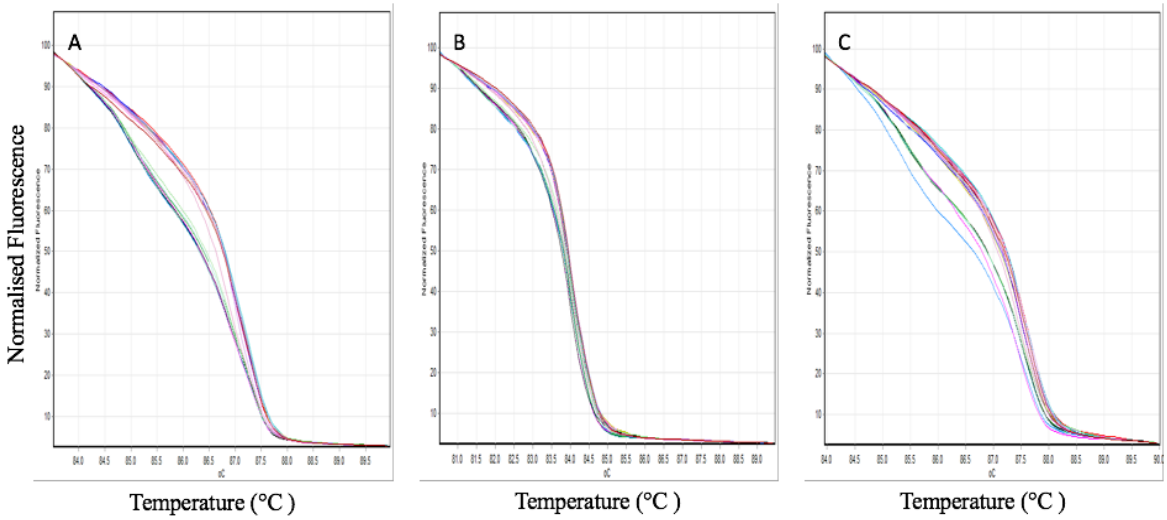


Figure 7 Normalised fluorescence graphs of HRM analysis from LpSSR9 within 20 genomic DNA within cv. Nui. A: AmpliTaq homemade master mix; B: Type-it HRM Kit; C, Roche Taq homemade master mix. Each line in the graph represents one sample.

4.1.2.6 LpSSR4 PCR product sequencing within cultivar Glencar

After HRM analysis of primer LpSSR4 and 29 individuals of cv. Glencar, nine clusters were classified by ScreenClust. PCR products from 12 individuals from the nine clusters were selected and sequencing plasmids were constructed successfully with a TOPO kit. In total, 22 plasmids were isolated from transformed cells and sent for sequencing. All sequencing results were aligned, and the sequences in the LpSSR4 region are shown in Figure 8. From the 22 alignment results, the LpSRR4 repeat motif is dinucleotide $(GT)_n$ and five different size repeat motif arrays (n=5, 6, 7, 9, and 10) exist in all sequences.

Apart from repeat motifs, two alleles of LpSSR4 also have three insertions identified at two sites and seven SNPs in the sequences flanking the SSR repeat motifs, and even one SNP in the repeat motifs, so the differences in these alleles are not only from different repeat motif numbers but also from SNPs and INDELs, and HRM analysis could detect all of them. Jones et al. (2001) presented the LpSSR4 DNA sequences in some pasture and turf grass species closely related to perennial ryegrass, and the SSR marker had a high degree of sequence conservation. The results here indicate that one allele of SSR4 in perennial ryegrass shares a 3-bp insertion (ACT) with Meadow fescue (*Festuca pratensis*) (Jones et al., 2001). It is

possible that gene flow happens between ryegrass and other grass species, such as Meadow fescue, due to there outcrossing nature (Kubik et al., 2001). Additionally, two alleles of LpSSR4 have the same repeat motifs, but have two SNPs in the flanking SSR repeat motifs. For gel electrophoresis, the difference in imperfect SSR loci can only be detected by length, while HRM analysis could tell not only length differences but also reveal additional mutations, such as SNPs in SSR loci (Lopez-Giraldez et al., 2007). This significant advantage could identify more polymorphisms in SSR markers, providing a deeper insight into population structures (Kayser et al., 2004).

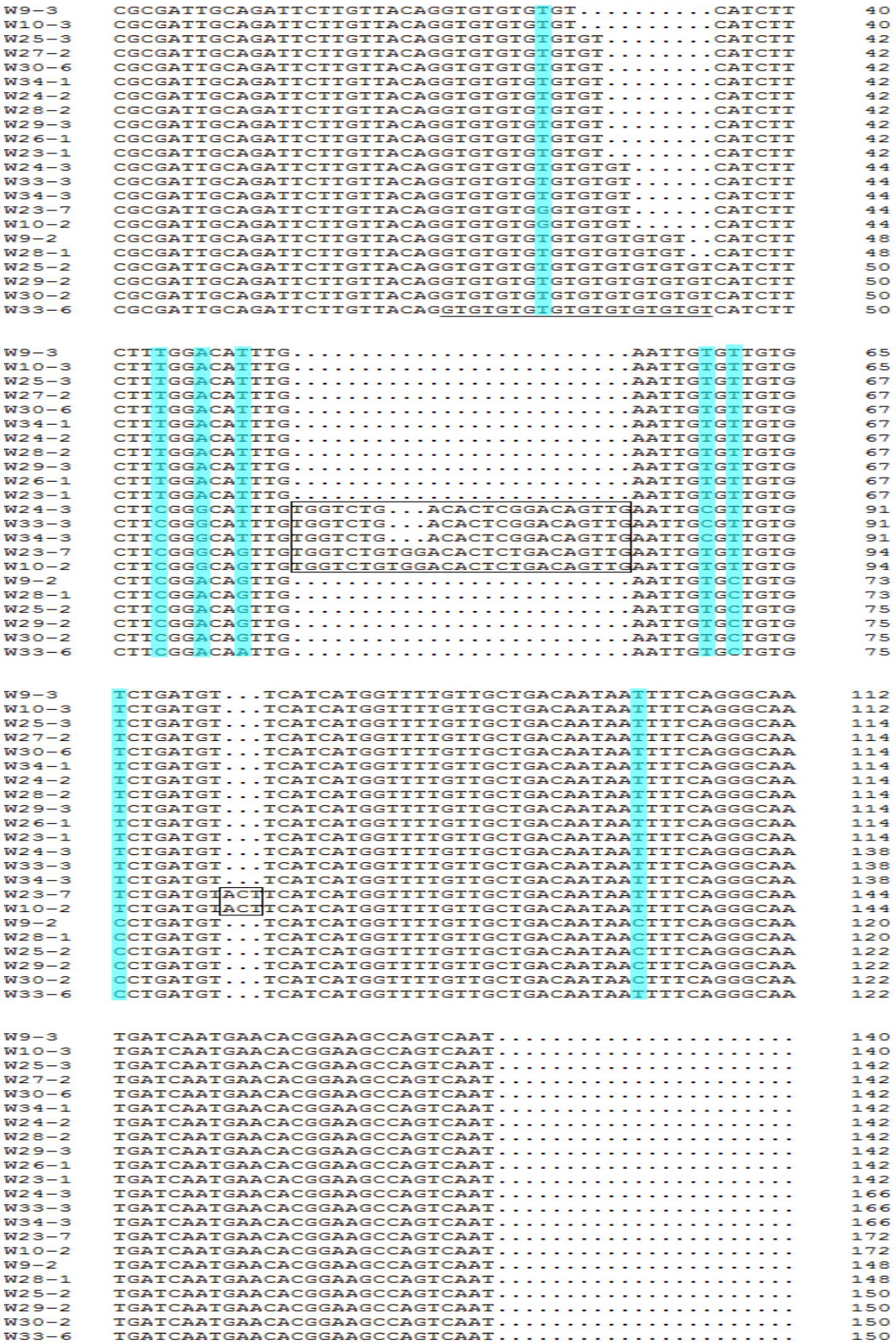


Figure 8 Alignment of LpSSR4 sequences from cv. Glencar. The motif repeat $(GT)_n$ is underlined, eight SNPs are highlighted in blue, and three insertions are boxed at two sites.

Appendix 4.2 GC plasmid sequence

The GC SNP site is highlighted.

TACGCAGCAGGTCTCATCAAGACGATCTACCCGAGTAACAATCTCCAGGAGATC
AAATACCTTCCCAAGAAGGTTAAAGATGCAGTCAAAAGATTCAGGACTAATTGC
ATCAAGAACACAGAGAAAGACATATTTCTCAAGATCAGAAGTACTATTCCAGTA
TGGACGATTCAAGGCTTGCTTCATAAACCAAGGCAAGTAATAGAGATTGGAGTC
TCTAAAAAGGTAGTTTCCTACTGAATCTAAGGCCATGCATGGAGTCTAAGATTCA
AATCGAGGATCTAACAGAACTCGCCGTGAAGACTGGCGAACAGTTCATACAGAG
TCTTTTACGACTCAATGACAAGAAGAAAATCTTCGTCAACATGGTGGAGCACGA
CACTCTGGTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGGGC
TATTGAGACTTTTCAACAAAGGATAATTTCGGGAAACCTCCTCGGATTCCATTGC
CCAGCTATCTGTCACTTCATCGAAAGGACAGTAGAAAAGGAAGGTGGCTCCTAC
AAATGCCATCATTGCGATAAAGGAAAGGCTATCATTCAAGATCTCTCTGCCGACA
GTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGAC
GTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGACATCTCCACTGACGTAA
GGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAG
TTCATTTCATTTGGAGAGGACACGCTCGAGATCACAAGTTTGTACAAAAAAGCA
GGCTTCATGAGACCAGAACGAAACCCCTTAGATCTTAACAATTTGCCCGATGAGT
ACTCTAGAGATGGCAAACAAGTCCTCGAAGACCATACCTCTTCATCCGGTTGCAG
GAAAAAGAAAAGCGGCGGGAAGGATGGAAAAGACGAGTGTGGGAAGGTCTACG
AGTGTAGATTTTGTTCCCTCAAGTTCTGCAAGTCTCAGGCTCTTGGGGGACACAT
GAACCGCCACCGCCAAGAGAGGGAAACGGAGACGCTGAACCAGGCTCGTCAAC
TGGTCTTTCGTTGTGATCATAACATTGCTGCACAAGGTGCCCCTCACTTAGGATG
CTGCCAAACAATAGGAACGGGGGGTTATCATCCCTCAGGAGACCCAACAGTGCC
TCTAAGATTCCCAAGATACTTCTCAGGTTCATCCTCAACTCACATGCCACCTCCCC
GCCACCGCCGCCGCCACCGCAACGACCATACCTATACCCTTCACCTACGAGGCCA
GTGTCATTTGGGTCATCACACTTCCCTCTCCAGCATGCAGTGAACGATTACTATG
TGGGCCACGTGATGAGTGGTGGCAGCCACGGACACTATGTTGGAGGAGAGAGCA
CAAGGAGTTACACGTGCATTGGTGCACCGGTGGGGCAAGGTGGCGGATTCGCTG
GTGGTAAGGAGGGGTCTGCAGTGCAGGAGGAAGGGTTGAGTACTT

Appendix 4.3 Trial of Guide-it mutation detection kit

The Guide-it Mutation Detection Kit (TaKaRa, Japan) was purchased to compare with the isolated CEL I nuclease. The PCR products were amplified from GC-CELI-FR373, FR648, and FR902 (Ref Table 4.1 for sequences) and G or C plasmid as template, using Terra PCR Direct Polymerase Mix and Buffer, supplied with the Guide-it Kit. The amplicons were then melted and hybridised to form heteroduplexes for a mismatched nucleotide for cleavage by the Guide-it Resolvase following the procedure outline in the manual. The products were separated on a 2% (w/v) agarose gel at 80 V for 50 min.

Three PCR products and their digestion products were performed by the Guide-it mutation detection kit (Figure 9 below). For the primer, GC-CELI-FR373, the PCR digestion products, the cleavage products sizes should be 157 bp and 216 bp, but only 216 bp cleavage product can be identified. For all three cleavage products, the 157 bp cleavage products cannot be detected from gel. The reason is that this product size (157 bp) is small and a number of cleavage products are also very small, so after 50 min electrophoresis, this DNA fragment cannot be observed. However, for FR648 and FR902, the cleavage products are at the wrong size, which means that non-specific PCR products were probably cut by the CEL I nuclease. Taken together, the PCR performed by the DNA polymerase from the Guide-it kit needs to be improved, and the Guide-it kit did not show a significantly better result than the CEL I isolated in the lab.

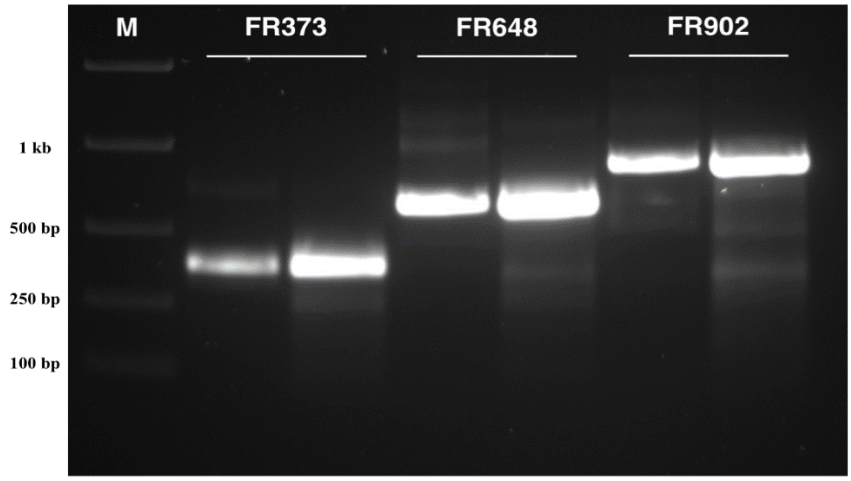


Figure 9 Three PCR products from primers GC-CELI-FR373, FR648, FR902 (left well of each primer) and their cleavage products (right well).

Appendix 4.4 Fluorescent dye selection

The GC-CELI-F1R1 primer (Ref Table 4.1 for seuqences) were used in the PCR to amplify the GC site with Bioline Taq DNA polymerase and Bioline 10x PCR buffer. The PCR products (5 µl) were digested by 1 µl of CEL I at 45°C for 30 min, and the digestion stopped by adding 5 µl of 0.225 M EDTA. A digestion reaction without CEL I was used as a control. Then 1.5 µl of EvaGreen dye (20x in water) or SYBR Green dye (20x in water) was added into the digestion products for HRM analysis and qPCR melt curve analysis, respectively. The HRM analysis was performed as follows: 60°C for 5 min, and then increase temperature to 95°C with a rate of 0.1°C/2sec; the qPCR melt curve analysis was performed as follows: 60°C for 5 min, and then increase temperature to 95°C at a rate of 1°C/5sec.

From the melting curve analysis with a different program, the results of EvaGreen could be identified products peaks, but in SYBR Green results, no peak is shown in qPCR melting curve. Therefore, EvaGreen was selected as a fluorescent dye in the following optimisation.

Appendix 4.5 Sequence of *LpCKX1* from cv. Nui

CGGTAACATCTATAGGGCGATTGATTTAGCGGCCGCGAATTCGCCCTTGC
CAGCCTTATCGCGCCTAAACGAGTCCCGCTAATGGCGATCGTTCACGTGT
TACTCGTCGCGCTAATCGCCGCGACGACGTCCCATGCACCGCGCGGCGTC
CATGGCGCCGCCGCCAAGACGACGTGGCCCGGCGACCTCGCCGCGCTCGC
CAAGGCCATCAGGCTCCGCACCGACCCGAACGCCACGGTGCCGGCCTCGA
CGGACTTCGGCAACATCACGTCAGCGCTCCCGGCGGCCGTGCTCCTCCCG
TCGTCCCCGGCGGACGTGGCCACGCTCCTCCGCGCCGCGCACTCCGACCC
CAAGTGGCCGTACACCATCTCCTTCCGCGGCCGCGGCCACTCCACCATGG
GCCAGGCCCTGGCCCCCGGCGGCGTGGTCGTCGACATGCCGTCCCTCGGC
GCCTCGTCGTCGCCCGCGGCGCCACGCATCAACGTTTCGGCGGACGGGCG
GTACGTGGACGCCGGCGGCGAGCAGATGTGGATCGACGTGCTGCGGGCGG
CTCTAGCGCGCGGCGTGGCGCCGCGGTCGTGGACGGACTACCTCCACCTC
ACCGTCGGCGGCACGCTCTCCAACGCCGGGATCAGCGG
CCAGACATTCCGGCACGGCCCGCAGATCTCCAACGTGCTGGAGCTCGACG
TCATCACCGGTACGTACGCGCGCACGTTACGGCCAGACATTCTTTTGCTA
GTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGCTTGTT
CCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCCGTGTA
CGCACAAGCTATGCTTTTTATGGTAATTTTTTTTACTATTTTCGGACGGT
ATAAAAATAGTCAAAAATTAATAATCATGACGAAAAATTATCATTGGCGA
GCACCTCACCAGCTTTACCGGCCGGCGCCGCCGTCTTTCATTAGATTTGC
TGCCGGCTCGGTTTACTGCGCGTTGACTTTCACATTTCCCCGCCGTTAAT
TAAATACTAGAATTATTACTCCTCTGCTCGAGTTGACAATTGGTGCAATT
GACGCACGCGTGCAGGCTACGGCGAGACGGTGACGTGCTCCAAGTCCCTC
AGCCCGGACCTCTTCAACGCGGTCCTGGGCGGGCTGGGCCAGTTCGGGGT
GATCGTCCGGGCCCGTATCGCGCTCGAGCCCGCGCCGGCCCGGGCCAAGT
GGACGCGGCTCGTCTACACCGACTTCGCCACCTTCTCCGCCGACCAGGAG
AAGCTCATCGCTCCCCGCCCGGACGGCTCGATCGGGCTGTTCAGTTACCT
CGAGGGCTCCGCCTTCGTCATCCACGGTCTCGCCGCCGCGCTGAAGAACTCTGGC
ACGTTCTTCTCCGACGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAG
AACGCCACCAACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACC
GCCGCGTCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGCTTCGAG
GAGGGGCTCTCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGGTGT
ACGGCGAGGAAATGACACTGGAGAAGCTCGGGCTGTGGCGCGTCCCGCACCCGT
GGCTCAACGTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGACCGCGGCGTCTT
CGGCGGCATCCTCCAGGGCACCGACATCGCTGGGCCTATGGTCATCTACCCGCTC
AACAAATCCAAGTACGCATTGATCATGCATGACATGATCATATATATCACGCGCT
TCATTTGCTCGTCGACGTCTGGTCTGAGTGTCTGACTGGCTCCGGTCGTTATTTCA
GGTGGGACGACAGCATGTCGGCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGG
TGTCGATGCTCTTCTCGTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGA
ACCAGAAGATCCTGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATC
TGGCGCATTACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGG
ACCGCTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAACTGCTCTCTCCAG
GACAGGACATCTTCAACTAGACTGAAGGGCGAATTCGTTTAAACCTGCAGGACT
AGTCCCTTTAGTGAGGCTAGTCTAGGCTGCCC

Appendix 4.6 Sequence of part *LpCKX2* from cv. Nui

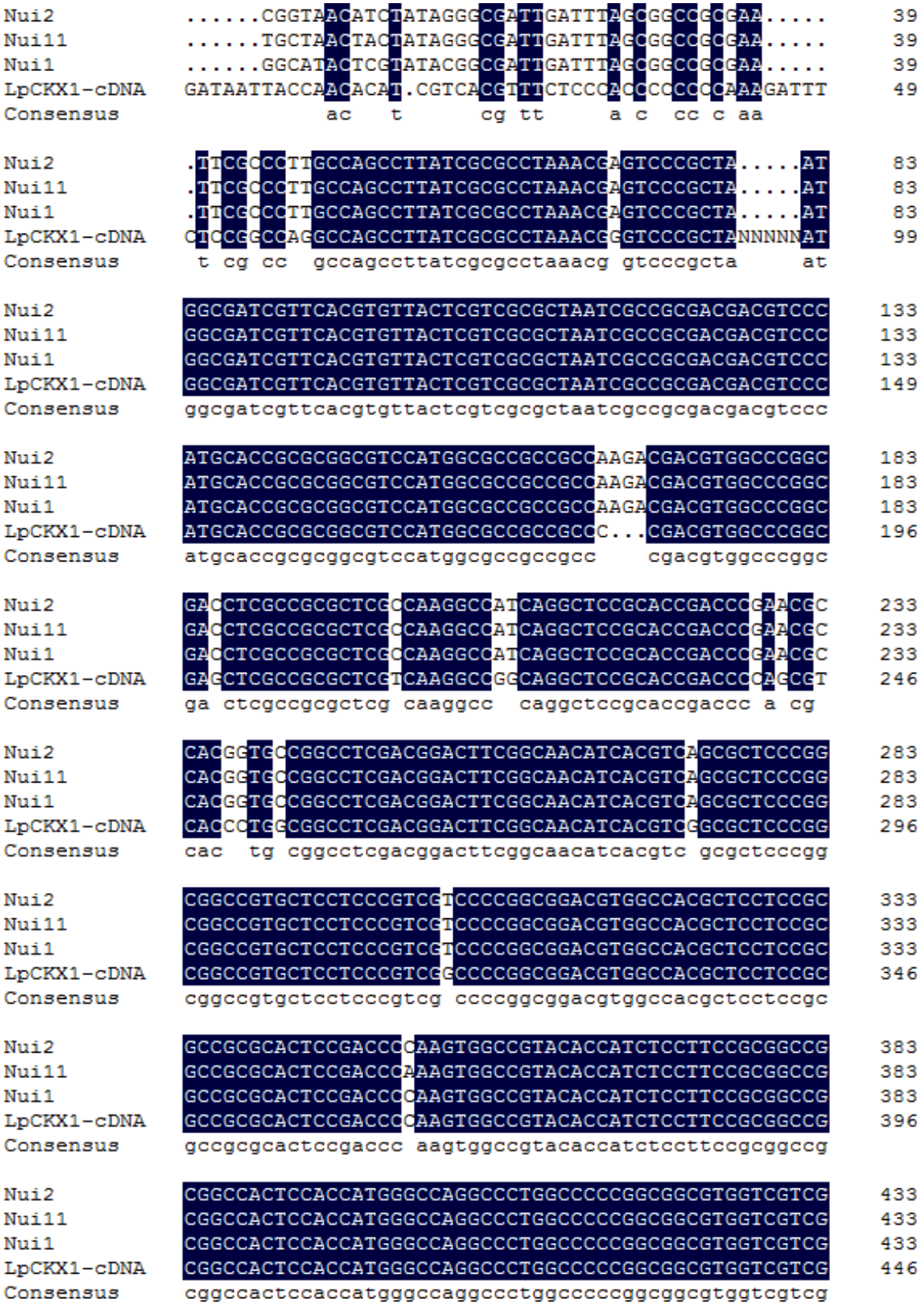
GACGCCGAGCGCAGCGTCGCTGCGCTCGCCAAGGAGGCAGGCGGCGTCCTGTAC
TGCCTCGAGGGAGCGCTGTACTACGGCGGCGCAGCCGGCAGCGAGCGCGACGTT
GATA(Exon 2)

AGGTAACCACACGACCTTTGTCGGTGCTGCTGCTAGCTATAGGGGACCACCAAGT
GAGATATGGATGGCAACTTTTTGCCGCCTATGACGCGGATGCATGGCGCCAAGTC
GCCAACTGTGCAAGCAGATCCGTGCAGCATCGGCGTCGTACATACTTGGCCTAGC
TTTGTCTGCCCTTGAAAATGCACCTCTTCTTGTCGCAATCTAACCCGTCCATATCC
GCAGTAATCATGGTTTAGGGATTAACTACACTACTACTTTACTATAAAAATTAAT
CTACTTGAAATAGTTTAGGGCAAGAGGTGTGCTTCGGTGCCCCCTCCCCCTCTCG
TCTTCCACCACCAGTAAAGTGGAAGGGATAAGTTGGCCCACGAAAGGGTATCGA
AAGATCTTGGCCGTTTAGAGGAGACCAAAGTTTGCAAAGGGGGACACCGTAGCA
ATGTTGTTAGGGCAAGAGGTGTGAAACGTTTTTCTTGCGTGGACGAATGGCAACT
CGCCACGTGGCCATGGAAAATATGCTAGAGTAGCCATCCACCGTTCAATTTTAAA
TCCGAGGGTTGTAGATTCGTTGTCGCGCTAGGATTCAAATGAGATAACACAATAT
AAAGTTTCCAATCTTTTAGGGAACACATCTAACTACTCGCTACTGGTTTTAATTA
GAGTCGACTCATTCATAAACCCTTTCCATTAAAGCAAAAAATCATGTGTATGCAA
ACTTAGGTTTGCTAAACCTAAACTTATACAAAAAAATCTACAAATCTACATCCAG
ACATTTTTTGTTGGATCAAGACTTCAATCAGTTCAATGCTCCCAACTAAGGTTTTG
ACCAAAGCACGGCCAACTTGCCCTCTCACATATCGTTGTTGCCAGTTTTGGACTG
TAGTGGCACATGCACATAAGAGTCAGACAAAGACACAGAAAAAAAAAAACTAAC
CCTATATGCAAACTAATAAGTGAGTAGGGACAGGATGTACAATCAATTTTGGGG
AGTCTCTTCATGGACTTAAAAAAAAACTCCTAG.CCAATTCCATGAACCCCCGGT
AAATGTTTTAGGGGGGAAATGTAAACCCGGGGTATTAAACTGCCTTTAGCTTCCA
GCTTTAATCCGGTGTGTGCAAATTTAGGTTTTTTATACCTAAACTTATACGAAAAT
TCTATAAATCTACATTTAGATTTGTTTTTGTTTGATCAAGACGTCAATGCTCTCAA
CTAAGGTTATGATGAAACCCGGCTAACTTCCCTCTCACGTATCTTATTTTTGTCTA
TTGGTCTGGAGCAACACATATATAAATACCAGATAAAATGCAGAAAGGACTAAT
CGTAAACTATGTGCAAATTGATATTCCCTTCGTTCCATAGTTCATGTCGTGGTTAA
GTCTAAATTTAAACTAGCTCCACAACAAGAATTATGGAACATGGGTTAAGTGAG
TAAGTGGCTACTAGGTACTAGGGAGTACTATATGAATATAGTCGAATAAATCAA
TTTTGTGTGTTTCTTCGTGAACTCTTTTTCGTATGTGTAATCCTTTTTGTACTTGCA
ATTCTAGCATGTGAAATTTCGCCAAAATTAAGTCTTCCATGTCGACATAGTGTCT
ATGTTTTCTAGCTAGCATCGATCAACAGCTAATACCTAGGACACGATCTAATGGT
GGGACAGCGACCGTGCATGGATGCATGCATGCACTCCGATATTCCTGGCTGGCTT
TACATATGCTGTACTAGTTCCACGCACACATGTCATACTGTACCCTCCACTTGTCC
AAGAAGTAACTTGCTCGATCACCTTTTGATAGTCTGCACTGTCTTTGGGGGAAAA
CATCACATGATCCCAAGCTCTCTGTCTCTGGTCCTACCTTCTTTGTCATCTCTTGT
AAGAAGATATAGCTGCAAAAGGTCATGACGATTCGTGTAGCCGTACATATTTTCA
GCTAGTAAACGTAAAAATATGCATATATGTGCCAACTTTACACTCTTGCATCCAT
GCCTACAAAATTTGACGTGACACGCATGTATCTTTGCTAAGTAATTTCTTTTTTTT
TGGCAATTTTTTTATAGTAATTTGTTAATTGTGCGC (Intron 2)

AGAAGCTGGATGTGCTGCTGCGGGAGTTGCGGTACGAGCGGGGCTTCGGGTTCG
TGCAGGACGTGTCGTACGTGGGGTTCCTGGACCGAGTGCGCGACGGTGAGCTCA
AGCTCCGCGCCACCGGTTTGTGGGAAGTGCCGCATCCCTGGCTCAACCTCTTCCT
CCCGCGCTCTCACGTCCTCGACTTCGCCGCCGGCGTCTTCCACGGCATCCTCCGC
CGAGGCACCACCGGCACCTCGAGGCCTGCCCTCATCTACCCCATGAACCGGAAC

AAGTGGGATGGCGAGGCTTCAGCGGTATTCCCGGAGGAGGAGGAGGTGTTCTAC
ACGGTGGGGATCCTGCGGTCGTCGGTGCCGGCGTCGTCGGACGGAGGCCAGCTG
CTGAGGCGCCTGGAGGAGCAGAACGAGGAGATCATGCGGTTCTGCGAGGATGCC
GGGATGGCGTGTGTGCAGTACCTGCCGTACTACGCCGGTCAGTCCGGCTGGGAG
AAGAAGCACTTCGGACCAAACAAGTGGGCCAGGTTCGTGGAGAGGAAGCGCAA
GTATGATCCCAAGGCGATCTTGTCCCGTGGACAGAGAATTTTCACATCCCCACTG
GCTTGA (Exon 3)

Appendix 4.7 Multiple sequence alignment of three full length of LpCKX1 cv. Nui1 and cDNA sequence of *LpCKX1* from the perennial ryegrass transcriptome database.



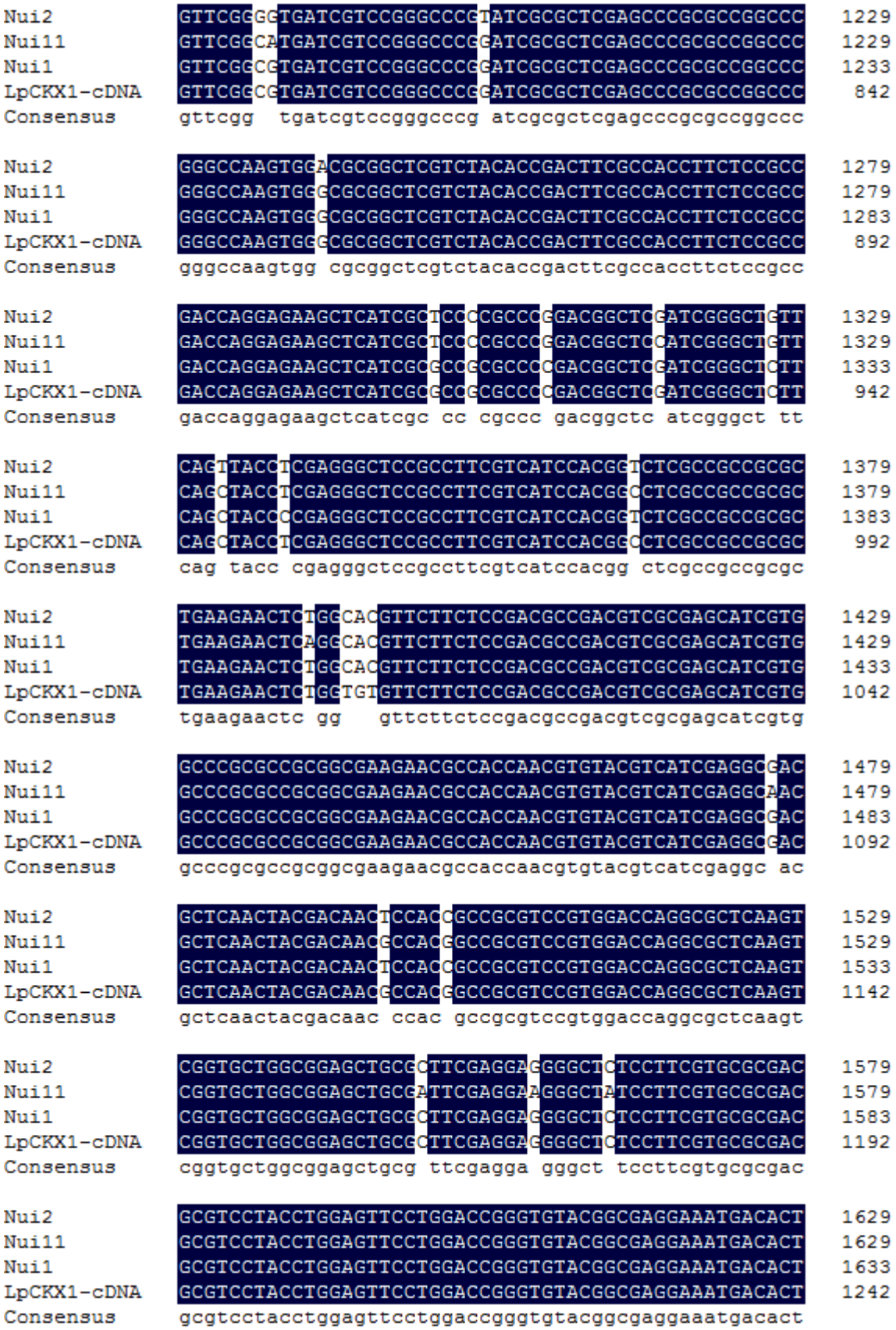
Appendix 4.7 Continued

```
Nui2          ACATGCCGTCCCTCGGCGCCTCGTCGTCGCCCGCGGCGCCACGCATCAAC   483
Nui11         ACATGCCGTCCCTCGGCGCCTCGTCGTCGCCCGCGGCGCCACGCATCAAC   483
Nui1          ACATGCCGTCCCTCGGCGCCTCGTCGTCGCCCGCGGCGCCACGCATCAAC   483
LpCKX1-cDNA   ACATGCCGTCCCTCGGCGCCTCGTCGTCGCCCGCGGCGCCACGCATCAAC   496
Consensus     acatgccgtccctcggcgcctcgtcgtcgcccgcggcgccacgcatcaac

Nui2          GTTTCGGCGGACGGGCGGTACGTGGACGCCGGCGGCGAGCAGATGTGGAT   533
Nui11         GTTTCGGCGGACGGGCGGTACGTGGACGCCGGCGGCGAGCAGATGTGGAT   533
Nui1          GTTTCGGCGGACGGGCGGTACGTGGACGCCGGCGGCGAGCAGATGTGGAT   533
LpCKX1-cDNA   GTTTCGGCGGACGGGCGGTACGTGGACGCCGGCGGCGAGCAGATGTGGAT   546
Consensus     gtttcggcggacgggcggtacgtggacgccggcggcgagcagatgtggat

Nui2          CGACGTGCTGCGGGCGGCTCTAGCGCGCGGCGTGGCGCCGCGGTCGTGGA   583
Nui11         CGACGTGCTGCGGGCGGCGCTGGCGCGCGGCGTGGCGCCGCGGTCGTGGA   583
Nui1          CGACGTGCTGCGGGCGGCGCTGGCGCGCGGCGTGGCGCCGCGGTCGTGGA   583
LpCKX1-cDNA   CGACGTGCTGCGGGCGGCGCTGGCGCGCGGCGTGGCGCCGCGGTCGTGGA   596
Consensus     cgacgtgctgcgggcggc ct gcgcgcggcgtggcgccgcggtcgtgga

Nui2          CGGACTACCTCCACCTCACCGTCGGCGGCACGCTCTCCAACGCCGGGATC   633
Nui11         CGGACTACCTCCACCTCACCGTCGGCGGCACGCTCTCCAACGCCGGGATC   633
Nui1          CGGACTACCTCCACCTCACCGTCGGCGGCACGCTCTCCAACGCCGGGATC   633
LpCKX1-cDNA   CGGACTACCTCCACCTCACCGTCGGCGGCACGCTCTCCAACGCCGGGATC   646
Consensus     cggactacctccacctcaccgtcggcggcacgctctccaacgccgggatc

Nui2          AGCGGCCAGACATTCCGGCACGGCCCGCAGATCTCCAACGTGCTGGAGCT   683
Nui11         AGCGGCCAGACATTCCGGCACGGCCCGCAGATCTCCAACGTGCTGGAGCT   683
Nui1          AGCGGCCAGACATTCCGGCACGGCCCGCAGATCTCCAACGTGCTGGAGCT   683
LpCKX1-cDNA   AGCGGCCAGACATTCCGGCACGGCCCGCAGATCTCCAACGTGCTGGAGCT   696
Consensus     agcggccagacattccggcacggcccgcagatctccaacgtgctggagct

Nui2          CGACGTCATCACCGGTACGTACGCGCGCACGTTACGGCCAGACATTCTTT   733
Nui11         CGACGTCATCACCGGTACGTACGCGCGCACGTTACGGCCAGACATTCTTT   733
Nui1          CGACGTCATCACCGGTACGTACGCGCGCACGTTACGGCCAGACATTCTTT   733
LpCKX1-cDNA   CGACGTCATCACTGNNNNNN..........................   716
Consensus     cgacgtcatcac g

Nui2          TGCTAGTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGC   783
Nui11         TGCTAGTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGC   783
Nui1          TGCTAGTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGC   783
LpCKX1-cDNA   ..................................................   716
Consensus

Nui2          TTGTTCCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCC   833
Nui11         TTGTTCCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCC   833
Nui1          TAGTTCCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCC   833
LpCKX1-cDNA   ..................................................   716
Consensus

Nui2          GTGTACGCACAAGCTATGCTTTTTATGGTAATTTTTTTTACTATTTTCGG   883
Nui11         GTGTACGCACAAGCTATGCTTTTTATGGTAATTTTTTTTACTATTTTCGG   883
Nui1          GTGTACGCACAAGCTATGCTTTTTATGGTAACTTTTTTTACTATTTTCGG   883
LpCKX1-cDNA   ..................................................   716
Consensus
```

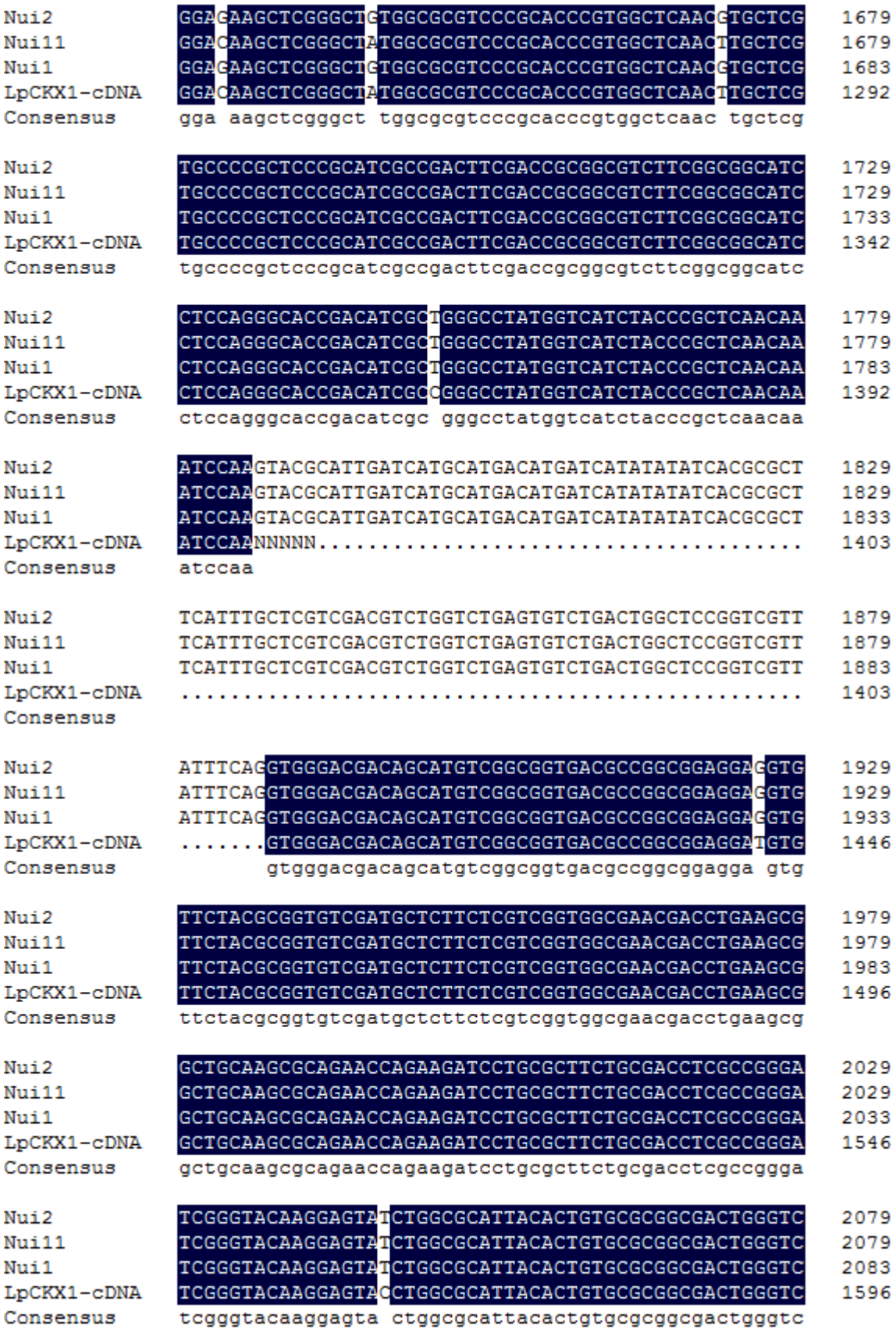
Appendix 4.7 Continued

```
Nui2          TGCTAGTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGC    783
Nui11         TGCTAGTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGC    783
Nui1          TGCTAGTTACGTGCCCCACACTGCGTTCTCTCCATTGATCACCAGCTAGC    783
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          TTGTTCCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCC    833
Nui11         TTGTTCCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCC    833
Nui1          TAGTTCCGGTGCAACTCGCCATTGATGAAGCACGGCAAATAGACCCGGCC    833
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          GTGTACGCACAAGCTATGCTTTTTATGGTAATTTTTTTTACTATTTTCGG    883
Nui11         GTGTACGCACAAGCTATGCTTTTTATGGTAATTTTTTTTACTATTTTCGG    883
Nui1          GTGTACGCACAAGCTATGCTTTTTATGGTAACTTTTTTTACTATTTTCGG    883
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          ACGGTATAAAAATAGTCAAAAATTAATAATCATGACGAAAAATTATCATT    933
Nui11         ACGGTATAAAAATAGTCAAAAATTAATAATCATGACGAAAAATTATCATT    933
Nui1          ATGGTATAAAAATAGTCAAAAAATAATAATCATGACGAAAAATTATCATT    933
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          GGCGAGC....ACCTCACCAGCTTTACCGGCCGGCGCCGCCGTCTTTCAT    979
Nui11         GGCGAGC....ACCTCACCAGCTTTACCGGCCGGCGCCGCCGTCTTTCAT    979
Nui1          GGCGAGCTAGCACCTCACCAGCTTTACCGGCCGGCGCCGCCGTCTTTCAT    983
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          TAGATTTGCTGCCGGCTCGGTTTACTGCGCGTTGACTTTCACATTTCCCC   1029
Nui11         TAGATTTGCTGCCGGCTCGGTTTACTGCGCGTTGACTTTCACATTTCCCC   1029
Nui1          TAGATTTGCTGCCGGCTCGGTTTACTACGCGTTGACTTTGACATTTCCCC   1033
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          GCCGTTAATTAAATACTAGAATTATTACTCCTCTGCTCGAGTTGACAATT   1079
Nui11         GCCGTTAATTAAATACTAGAATTATTACTCCTCTGCTCGAGTTGACAATT   1079
Nui1          GCCGTTAATTAAATACTAGAATTATTACTCCTCTGCTCGAGTTGACAATT   1083
LpCKX1-cDNA   ..................................................    716
Consensus

Nui2          GGTGCAATTGACGCACGCGTGCAGGCTACGGCGAGACGGTGACGTGCTCC   1129
Nui11         GGTGCAATTGACGCACGCGTGCAGGCTACGGCGAGACGGTGACGTGCTCC   1129
Nui1          GGTGCAATTAACGCACGCGTGCAGGCTACGGCGAGACGGTGACGTGCTCC   1133
LpCKX1-cDNA   ......................GCTACGGCGAGACGGTGACGTGCTCC    742
Consensus                           gctacggcgagacggtgacgtgctcc

Nui2          AAGTCCCTCAGCCCGGACCTCTTCAACGCGGTCCTGGGCGGGCTGGGCCA   1179
Nui11         AAGTCCCTCAGCCCGGACCTCTTCAACGCGGTCCTGGGCGGGCTGGGCCA   1179
Nui1          AAGTCCCTCAGCCCGGACCTCTTCAACGCGGTCCTGGGCGGGCTGGGCCA   1183
LpCKX1-cDNA   AAGTCCCTCAGCCCGGACCTCTTCAACGCGGTCCTTGGCGGGCTGGGCCA    792
Consensus     aagtccctcagcccggacctcttcaacgcggtcct ggcgggctgggcca
```

Appendix 4.7 Continued



Appendix 4.7 Continued



Appendix 4.7 Continued

```
Nui2          CGGCATTTCGGCGGCAAGTGGGACCGCTTCGTCCAGATGAAGGACAAGTA   2129
Nui11         CGGCATTTCGGCGGCAAGTGGGACCGCTTCGTCCAGATGAAGGACAAGTA   2129
Nui1          CGGCATTTCGGCGGCAAGTGGGACCGCTTCGTCCAGATGAAGGACAAGTA   2133
LpCKX1-cDNA   CGGCATTTCGGCGGCAAGTGGGACCGCTTCGTCCAGATGAAGGACAAGTA   1646
Consensus     cggcatttcggcggcaagtgggaccgcttcgtccagatgaaggacaagta

Nui2          CGACCCCAAGAAACTGCTCTCTCCAGGACAGGACATCTTCAACTAG....   2175
Nui11         CGACCCCAAGAAACTGCTCTCTCCAGGACAGGACATCTTCAACTAG....   2175
Nui1          CGACCCCAAGAAACTGCTCTCTCCAGGACAGGACATCTTCAACTAG....   2179
LpCKX1-cDNA   CGACCCCAAGAAACTGCTCTCTCCAGGACAGGACATCTTCAACTAGNNNN   1696
Consensus     cgaccccaagaaactgctctctccaggacaggacatcttcaactag

Nui2          ..ACTGAAGGGCGAATTCGTTTAAACCTGCAGGA.CTAGTCCCTTTAGTG   2222
Nui11         ..ACTGAAGGGCGAATTCGTTTAAACCTGCAGGA.CTAGTCCCTTTAGTG   2222
Nui1          ..ACTGAAGGGCGAATTCGTTTAAACCTGCAGGA.CTAGTCCCTTTAGTG   2226
LpCKX1-cDNA   NNACTGCAGATAATTGTTGCTAGCTCCATCAAAAGTTAGTGGATTAACAT   1746
Consensus       actg ag       t g t    cc  ca  a  tagt   tt a

Nui2          AGGCTAGTCTAGGCTGCCC.............................   2241
Nui11         AGGCTATCTAGCGCCG................................   2238
Nui1          AGGT..ATCTAGGCCACCCCA...........................   2245
LpCKX1-cDNA   GGACAGACACAAAGCAGTTCAGTAAACATTTGGGTTTTAGTACTCTCTTG   1796
Consensus      g

Nui2          ................................................   2241
Nui11         ................................................   2238
Nui1          ................................................   2245
LpCKX1-cDNA   CAAAGGCAAATAGGTGGTAAACGTTAGTAATAATGTGTTCATCTAATAAT   1846
Consensus

Nui2          ................................................   2241
Nui11         ................................................   2238
Nui1          ................................................   2245
LpCKX1-cDNA   ATGTTCTATCCACCTCAAATCTCCATATGTACCCATCGAATTATAGGAGT   1896
Consensus

Nui2          ...............................                  2241
Nui11         ...............................                  2238
Nui1          ...............................                  2245
LpCKX1-cDNA   ATATATAAAGTAATAAAAAAAGTGTGCATATC                   1928
Consensus
```

Appendix 4.8 Multiple sequence alignment of sequencing results of primer LpCKX1-F10R10 amplified from cv. Nui.

```
d2b     CGGTCCTGGGCGGGCTGGGCCAGTTCGGCGTGATCGTCCGGGCCCGGATC     50
d4      CGGTCCTGGGCGGGCTGGGCCAGTTCGGCGTGATCGTCCGGGCCCGGATC     50
d3b     CGGTCCTGGGCGGGCTGGGCCAGTTCGGCGTGATCGTCCGGGCCCGGATC     50
d3a     CGGTCCTGGGCGGGCTGGGCCAGTTCGGCGTGATCGTCCGGGCCCGGATC     50
d9b     CGGTCCTGGGCGGGCTGGGCCAGTTCGGGGTGATCGTCCGGGCCCGGATC     50
d15b    CGGTCCTGGGCGGGCTGGGCCAGTTCGGCATGATCGTCCGGGCCCGGATC     50
d15a    CGGTCCTGGGCGGGCTGGGCCAGTTCGGCATGATCGTCCGGGCCCGGATC     50
d9a     CGGTCCTGGGCGGGCTGGGCCAGTTCGGCGTGATCGTCCGGGCCCGGATC     50
d2a     CGGTCCTGGGCGGGCTGGGCCAGTTCGGGGTGATCGTCCGGGCCCGTATC     50
d12     CGGTCCTGGGCGGGCTGGGCCAGTTCGGCGTGATCGTCCGGGCCCGGATC     50
Consenscggtcctgggcgggctgggccagttcgg  tgatcgtccgggcccg atc

d2b     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d4      GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d3b     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d3a     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d9b     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d15b    GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d15a    GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d9a     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
d2a     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGACGCGGCTCGTCTACAC     100
d12     GCGCTCGAGCCCGCGCCGGCCCGGGCCAAGTGGGCGCGGCTCGTCTACAC     100
Consensgcgctcgagcccgcgccggcccgggccaagtgg cgcggctcgtctacac

d2b     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCGCCGCGCC     150
d4      CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCGCCGCGCC     150
d3b     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCGCCGCGCC     150
d3a     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCGCCGCGCC     150
d9b     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCTCCGCGCC     150
d15b    CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCTCCCCGCC     150
d15a    CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCTCCCCGCC     150
d9a     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCTCCCCGCC     150
d2a     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCTCCCCGCC     150
d12     CGACTTCGCCACCTTCTCCGCCGACCAGGAGAAGCTCATCGCGCCGCGCC     150
Consenscgacttcgccaccttctccgccgaccaggagaagctcatcgc cc cgcc

d2b     CCGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d4      CCGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d3b     CCGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d3a     CCGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d9b     CGGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d15b    CGGACGGCTCCATCGGGCTGTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d15a    CGGACGGCTCCATCGGGCTGTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d9a     CCGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
d2a     CGGACGGCTCGATCGGGCTGTTCAGTTACCTCGAGGGCTCCGCCTTCGTC     200
d12     CCGACGGCTCGATCGGGCTCTTCAGCTACCTCGAGGGCTCCGCCTTCGTC     200
Consenscsgacggctc atcgggct ttcag tacctcgagggctccgccttcgtc
```

Appendix 4.8 Continued

```
d2b     ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
d4      ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
d3b     ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
d3a     ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCAGGCACGTTCTTCTCCGA    250
d9b     ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCAGGCACGTTCTTCTCCGA    250
d15b    ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
d15a    ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCAGGCACGTTCTTCTCCGA    250
d9a     ATCCACGGCCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
d2a     ATCCACGGTCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
d12     ATCCACGGTCTCGCCGCCGCGCTGAAGAACTCTGGCACGTTCTTCTCCGA    250
Consensatccacgg ctcgccgccgcgctgaagaactc ggcacgttcttctccga

d2b     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d4      CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d3b     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d3a     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d9b     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d15b    CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d15a    CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d9a     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d2a     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
d12     CGCCGACGTCGCGAGCATCGTGGCCCGCGCCGCGGCGAAGAACGCCACCA    300
Consenscgccgacgtcgcgagcatcgtggcccgcgccgcggcgaagaacgccacca

d2b     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACGCCACGGCCGCG    350
d4      ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACGCCACGGCCGCG    350
d3b     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACGCCACGGCCGCG    350
d3a     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACGCCACGGCCGCG    350
d9b     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACGGCCGCG    350
d15b    ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACCGCCGCG    350
d15a    ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACGGCCGCG    350
d9a     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACCGCCGCG    350
d2a     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACCGCCGCG    350
d12     ACGTGTACGTCATCGAGGCGACGCTCAACTACGACAACTCCACCGCCGCG    350
Consensacgtgtacgtcatcgaggcgacgctcaactacgacaac ccac gccgcg

d2b     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d4      TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d3b     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d3a     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d9b     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d15b    TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d15a    TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d9a     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGATTCGAGGA    400
d2a     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGCTTCGAGGA    400
d12     TCCGTGGACCAGGCGCTCAAGTCGGTGCTGGCGGAGCTGCGCTTCGAGGA    400
Consenstccgtggaccaggcgctcaagtcggtgctggcggagctgcg ttcgagga
```

Appendix 4.8 Continued

```
d2b     AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d4      AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d3b     AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d3a     AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d9b     AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d15b    AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d15a    AGGGCTATCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d9a     AGGGCTCTCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d2a     GGGGCTCTCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
d12     GGGGCTCTCCTTCGTGCGCGACGCGTCCTACCTGGAGTTCCTGGACCGGG    450
Consensugggct tccttcgtgcgcgacgcgtcctacctggagttcctggaccggg

d2b     TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d4      TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d3b     TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d3a     TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d9b     TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d15b    TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d15a    TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d9a     TGTACGGCGAGGAAATGACACTGGACAAGCTCGGGCTATGGCGCGTCCCG    500
d2a     TGTACGGCGAGGAAATGACACTGGAGAAGCTCGGGCTGTGGCGCGTCCCG    500
d12     TGTACGGCGAGGAAATGACACTGGAGAAGCTCGGGCTGTGGCGCGTCCCG    500
Consenstgtacggcgaggaaatgacactgga aagctcgggct tggcgcgtcccg

d2b     CACCCGTGGCTCAACTTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d4      CACCCGTGGCTCAACTTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d3b     CACCCGTGGCTCAACTTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d3a     CACCCGTGGCTCAACTTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d9b     CACCCGTGGCTCAACGTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d15b    CACCCGTGGCTCAACTTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d15a    CACCCGTGGCTCAACGTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d9a     CACCCGTGGCTCAACTTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d2a     CACCCGTGGCTCAACGTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
d12     CACCCGTGGCTCAACGTGCTCGTGCCCCGCTCCCGCATCGCCGACTTCGA    550
Consenscacccgtggctcaac tgctcgtgccccgctcccgcatcgccgacttcga

d2b     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d4      CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d3b     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d3a     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d9b     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d15b    CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d15a    CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d9a     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d2a     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
d12     CCGCGGCGTCTTCGGCGGCATCCTCCAG                          578
Consensccgcggcgtcttcggcggcatcctccag
```

Appendix 4.9 Multiple sequence alignment of sequencing result of intron 2 to exon 3 in *LpCKX1* amplified by LpCKX1-F13R15 from 11 individuals from cv. Nui.

```
D9-2     TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D2-2     TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D15-2    TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D20-2    TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D1-2     TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D12      TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D14      TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D1-1     TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D9-1     TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D2-1     TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D15-1    TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D20-1    TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D3       TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D4       TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D6       TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
D10      TATGGTCATCTACCCGCTCAACAAATCCAAGTACGCATTGATCATGCATG     50
Consensutatggtcatctacccgctcaacaaatccaagtacgcattgatcatgcatg

D9-2     ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D2-2     ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D15-2    ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D20-2    ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D1-2     ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D12      ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D14      ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGTCTGGTCTG     100
D1-1     ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D9-1     ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D2-1     ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D15-1    ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D20-1    ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D3       ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D4       ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D6       ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
D10      ACATGATCATATATATCACGCGCTTCATTTGCTCGTCGACGACTGGTCTG     100
Consensuacatgatcatatatatcacgcgcttcatttgctcgtcgacg ctggtctg

D9-2     AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D2-2     AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D15-2    AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D20-2    AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D1-2     AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D12      AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D14      AGTGTCTGACTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     150
D1-1     A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D9-1     A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D2-1     A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D15-1    A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D20-1    A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D3       A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D4       A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D6       A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
D10      A.......CTGGCTCCGGTCGTTATTTCAGGTGGGACGACAGCATGTCG     142
Consensua        ctggctccggtcgttatttcaggtgggacgacagcatgtcg
```

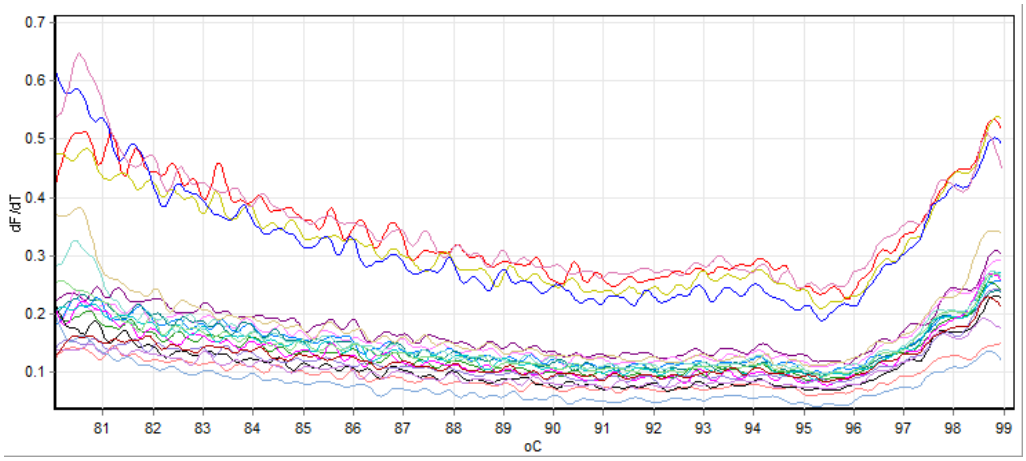
Appendix 4.9 Continued

```
D9-2      GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D2-2      GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D15-2     GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D20-2     GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D1-2      GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D12       GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D14       GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    200
D1-1      GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D9-1      GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D2-1      GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D15-1     GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D20-1     GCGGTGACGCCGGCGGAGGAGGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D3        GCGGTGACGCCGGCGGAGGATGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D4        GCGGTGACGCCGGCGGAGGATGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D6        GCGGTGACGCCGGCGGAGGATGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
D10       GCGGTGACGCCGGCGGAGGATGTGTTCTACGCGGTGTCGATGCTCTTCTC    192
Consensugcggtgacgccggcggagga gtgttctacgcggtgtcgatgctcttctc

D9-2      GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D2-2      GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D15-2     GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D20-2     GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D1-2      GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D12       GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D14       GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    250
D1-1      GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D9-1      GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D2-1      GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D15-1     GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D20-1     GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D3        GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D4        GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D6        GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
D10       GTCGGTGGCGAACGACCTGAAGCGGCTGCAAGCGCAGAACCAGAAGATCC    242
Consensugtcggtggcgaacgacctgaagcggctgcaagcgcagaaccagaagatcc

D9-2      TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D2-2      TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D15-2     TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D20-2     TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D1-2      TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D12       TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D14       TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    300
D1-1      TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D9-1      TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D2-1      TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D15-1     TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D20-1     TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D3        TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D4        TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D6        TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
D10       TGCGCTTCTGCGACCTCGCCGGGATCGGGTACAAGGAGTATCTGGCGCAT    292
Consensutgcgcttctgcgacctcgccgggatcgggtacaaggagtatctggcgcat
```

Appendix 4.9 Continued

```
D9-2     TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D2-2     TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D15-2    TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D20-2    TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D1-2     TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D12      TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D14      TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    350
D1-1     TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D9-1     TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D2-1     TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D15-1    TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D20-1    TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D3       TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D4       TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D6       TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
D10      TACACTGTGCGCGGCGACTGGGTCCGGCATTTCGGCGGCAAGTGGGACCG    342
Consensutacactgtgcgcggcgactgggtccggcatttcggcggcaagtgggaccg

D9-2     CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D2-2     CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D15-2    CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D20-2    CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D1-2     CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D12      CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D14      CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 387
D1-1     CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D9-1     CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D2-1     CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D15-1    CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D20-1    CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D3       CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D4       CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D6       CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
D10      CTTCGTCCAGATGAAGGACAAGTACGACCCCAAGAAA                 379
Consensucttcgtccagatgaaggacaagtacgaccccaagaaa
```

Appendix 4.10 The melting curve of HRM analysis from the trial 2 of the new method for CEL I product detection. Each colour line presents one CEL I digestion products. When the reactions were heated to 99°C, the DNA fragments are still not totally denatured.



Appendix 5.1 All reaction components and the cycling program for HRM master mixes

HRM PCR Master Mix for three commercial kits (Type-it, Biotium, and Brilliant)

Component	Volume per 15- µL reaction (µL)	Final concentration
2x HRM PCR Master Mix	7.5	1x
Template DNA [a]	1	
5 µM Primer Mix [b]	1	0.33 µM
PCR-grade water	Add to 15 µL	

a, the template DNA concentration is about 50 pg/µL.
b, 5 µM primer mix consists of 5 µM forward primer and 5 µM reverse primer.

Cycling protocol for HRM analysis with Type-it HRM PCR Master Mix

Stage	Step	Temp.	Time
Initial PCR step	Denaturation/ Activation	95°C	5 min
Cycling, 40 cycles	Denature	95°C	10 sec
	Annealing/Extension	55°C	30 sec
HRM	Fluorescence data acquisition	70-90°C, 0.1°C increments	2 sec

Cycling protocol for HRM analysis with Biotium Fast EvaGreen qPCR Master Mix

Stage	Step	Temp.	Time
Initial PCR step	Denaturation/ Activation	95°C	2 min
Cycling, 40 cycles	Denature	95°C	5 sec
	Annealing/Extension	60°C	30 sec
HRM	Fluorescence data acquisition	70-90°C, 0.1°C increments	2 sec

Cycling protocol for HRM analysis with Brilliant Agilent HRM Ultra-Fast Loci Master Mix

Stage	Step	Temp.	Time
Initial PCR step	Denaturation/ Activation	95°C	2 min
Cycling, 40 cycles	Denature	95°C	5 sec
	Annealing/Extension	60°C	30 sec
HRM	Fluorescence data acquisition	70-90°C, 0.1°C increments	2 sec

AmpliTaq Gold 360 DNA Polymerase

Component	Volume per 15- µL reaction (µL)	Final concentration
AmpliTaq Gold 360 Buffer, 10x	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
25 mM MgCl2[a]	1.8	3 mM
AmpliTaq Gold 360 DNA Polymerase	0.075	0.375 Units
Template DNA	1	
5 µM Primer Mix [b]	1	0.33 µM
EvaGreen	0.75	1x
360 GC Enhancer [c]	0.3	2%
PCR-grade water	Make up to 15 µL	

a 25 mM $MgCl_2$ supplied by AmpliTaq Gold 360 DNA Polymerase kit
b A 5 µM primer mix consists of 5 µM forward primer and 5 µM reverse primer.
c 360 GC Enhancer supplied by AmpliTaq Gold 360 DNA polymerase kit

Cycling protocol for HRM analysis with AmpliTaq Gold 360 DNA Polymerase

Stage	Step	Temp.	Time
Initial PCR activation	Activation of AmpliTaq Gold 360 DNA Polymerase	95°C	10 min
Cycling, 40 cycles	Denature	95°C	30 sec
	Annealing/Extension	55°C	1 min
HRM	Fluorescence data acquisition	70-90°C, 0.1°C increments	2 sec

Roche Taq DNA polymerase

Component	Volume per 15- µL reaction (µL)	Final concentration
Roche Taq PCR reaction buffer, 10x	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
25 mM $MgCl_2$	1.8	3 mM
Taq DNA Polymerase	0.075	0.375 Units
Template DNA	1	
5 µM Primer Mix	1	0.33 µM
EvaGreen	0.75	1x
PCR-grade water	Make up to 15 µL	

Cycling protocol for HRM analysis with Roche Taq DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Initial denaturation	95°C	2 min
Cycling, 40 cycles	Denaturation	95°C	15 sec
	Annealing	58°C	20 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	75-90°C, 0.1°C increments	2 sec

Bioline Taq DNA Polymerase

Component	Volume per 15- µL reaction (µL)	Final concentration
Bioline Taq PCR reaction buffer, 10x	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
50 mM $MgCl_2$	0.5	1.6 mM
Taq DNA Polymerase	0.2	1 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	1	0.33 µM
EvaGreen	0.75	1x
PCR-grade water	Make up to 15 µL	

Cycling protocol for HRM analysis with Bioline Taq DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Initial denaturation	95°C	10min
Cycling, 40 cycles	Denaturation	95°C	10 sec
	Annealing	58°C	15 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	75-90°C, 0.1°C increments	2 sec

HotStarTaq *Plus* DNA Polymerase

Component	Volume per 15- µL reaction (µL)	Final concentration
HotStar PCR buffer, 10x	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
HotStarTaq *Plus* DNA Polymerase	0.075	0.375 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	1	0.33 µM
EvaGreen	0.75	1x
PCR-grade water	Make up to 15 µL	

Cycling protocol for HRM analysis with HotStarTaq *Plus* DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Initial denaturation	95°C	5 min
Cycling, 40 cycles	Denaturation	95°C	15 sec
	Annealing	55°C	20 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	75-90°C, 0.1°C increments	2 sec

FastStart Taq DNA Polymerase (Roche)

Component	Volume per 15- µL reaction (µL)	Final concentration
FastStart Taq PCR reaction buffer, 10x	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
FastStart Taq DNA Polymerase	0.12	0.6 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	1	0.33 µM
EvaGreen	0.75	1x
PCR-grade water	Make up to 15 µL	

Cycling protocol for HRM analysis with FastStart Taq DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Denaturation/ Activation	95°C	4 min
Cycling, 40 cycles	Denaturation	95°C	20 sec
	Annealing	58°C	20 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	75-90°C, 0.1°C increments	2 sec

GoTaq DNA Polymerase

Component	Volume per 15- µL reaction (µL)	Final concentration
GoTaq reaction buffer, 5x	3	1x
20 mM dNTP mix	0.15	0.2 mM each
GoTaq DNA Polymerase	0.075	0.375 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	1	0.33 µM
EvaGreen	0.75	1x
PCR-grade water	Make up to 15 µL	

Cycling protocol for HRM analysis with GoTaq DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Denaturation/ Activation	95°C	2 min
Cycling, 40 cycles	Denaturation	95°C	20 sec
	Annealing	58°C	20 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	75-90°C, 0.1°C increments	2 sec

DyNAzyme II DNA Polymerase

Component	Volume per 15- µL reaction (µL)	Final concentration
10x DyNAzyme buffer	1.5	1x
20 mM dNTP mix	0.15	0.2 mM each
DyNAzyme II DNA Polymerase	0.075	0.375 Units per 15 reaction
Template DNA	1	
5 µM Primer Mix	1	0.33 µM
EvaGreen	0.75	1x
PCR-grade water	Make up to 15 µL	

Cycling protocol for HRM analysis with DyNAzyme II DNA Polymerase

Stage	Step	Temp.	Time
Initial Denaturation	Denaturation	95°C	2 min
Cycling, 40 cycles	Denaturation	95°C	20 sec
	Annealing	58°C	20 sec
	Extension	72	20 sec
HRM	Fluorescence data acquisition	75-90°C, 0.1°C increments	2 sec

Appendix 5.2

	LpSH1-HRM-F2R2	LpSH1-HRM-F7R5		LpSH1-HRM-F2R2	LpSH1-HRM-F7R5		LpSH1-HRM-F2R2	LpSH1-HRM-F7R5		LpSH1-HRM-F2R2	LpSH1-HRM-F7R5		LpSH1-HRM-F2R2	LpSH1-HRM-F7R5		LpSH1-HRM-F2R2	LpSH1-HRM-F7R5		LpSH1-HRM-F2R2	LpSH1-HRM-F7R5
1	A	A	101	A	A	201	A	B	301	A	A	401	A		501	A	B	601	D	C
2	A	A	102			202	A	B	302	A	B	402	A	D	502	A	A	602	A	B
3	B	C	103	A	A	203	A	B	303	A	A	403	C	403	503	A	B	603	A	D
4	A	B	104	A	B	204	A	B	304	D	C	404	A	D	504	504	D	604	A	B
5	A	B	105	A	B	205	A	B	305	C	C	405	B	C	505	A	B	605	A	A
6	C	D	106	A	A	206	A	B	306	D	C	406	A	B	506	A	B	606	A	A
7	A	B	107	A	A	207	A	B	307	A	B	407	A	A	507	A	B	607	A	A
8	A	B	108	A?	A	208	A	B	308	D	C	408	C	403	508	A	B	608	A	A
9	A	B	109			209	A	B	309	A	B	409	A	B	509	504	C	609	A	A
10	C	D	110	A	A	210	A	B	310	A	B	410	D		510	C	C	610	A	B
11	A	B	111	A	B	211	A	B	311	A	A	411	D	D	511	A	B	611	A	B
12	D	C	112	A	A	212	A	B	312	A	A	412	D	D	512	A	B	612	B	C
13	A	B	113			213	A	B	313	A	A	413	A	D	513	A	A	613	A	C
14	B	C	114	A	B	214	D	D	314	A	C	414	A	403	514	A	B	614	C	D
15	B	C	115			215	D	D	315	A	B	415	A	A	515	A	B	615	A	A
16			116			216	A	B	316	A	B	416	A	A	516	A	B	616	A	A
17	A	B	117	A	B	217	A	B	317	A	B	417	C	D	517	D	C	617	A	A
18	A	B	118			218	A	B	318	D	C	418	C	D	518	A	A	618	C	D
19	A	B	119			219	A	B	319	A	B	419	A	D	519	A	B	619	A	C
20	A	A	120			220	D	D	320	A	A	420	A	B	520	C	C	620	A	A
21	21	C	121	A	A	221	A	B	321	C	C	421	A	B	521	A	B	621	A	A
22	B	C	122	A	B	222	A	B	322	A	B	422	A	B	522	A	A	622	A	A
23	C	C	123	A	A	223	A	B	323	A	A	423	A	B	523	A	B	623	A	A
24	A	B	124	A	A	224	A	B	324	A	B	424	A	B	524	C	D	624	A	A
25	B	C	125	A	A	225	A	B	325	A	B	425	A	B	525	A	B	625	A	B
26	C	C	126	A	A	226	A	B	326	A	B	426	A	B	526	A	B	626	A	B
27	A	A	127	A		227	A	B	327	A	B	427	A	B	527	A	B	627	A	D
28	B	D	128	A	B	228	A	B	328	D	C	428	A	B	528	A	B	628	A	A
29	B	C	129	A	A	229	A	B	329	A	B	429	A	B	529	A	B	629	A	A
30			130	A	A	230	A	B	330	A	A	430	A	B	530	A	B	630	A	D
31	31	C	131	A	A	231	A	B	331	A	A	431			531	A	B	631	B	C
32	C	D	132	A	A	232	A	B	332	A	A	432	A	B	532	A	B	632	A	A
33	B	C	133			233	A	B	333	A	A	433	A	433	533	A	B	633	A	B
34	B	C	134	A	A	234	A	B	334	A	A	434	A	B	534	A	B	634	A	A
35	B	C	135			235	A	B	335	A	A	435	A	B	535	C	D	635	A	A
36	36	C	136	A	A	236	A	A	336	A	A	436	A	B	536	A	B	636	A	A
37		C	137			237	A	B	337	A	B	437	A	B	537	A	B	637	A	B
38	B	D				238	A	B	338	A	A	438	A	B	538	A	B	638	A	D
39	B	C				239	A	B	339	A	B	439	A	A	539	A	B	639	A	A
40	B	C				240	A	B	340	A	A	440	A	B	540	A	B	640	A	D
41	A	B	141	A	B	241	A	B	341	A	B	441	A	B	541	A	B	641	A	A
42	A	D	142	A	B	242	A	B	342	A	B	442	A	B	542	A	B	642	A	A
43	A	A	143	A	B	243	A	B	343	A	B	443	A	B	543	A	B	643	A	D
44	A	A	144	A	B	244	B	C	344	A	B	444	A	B	544	A	B	644	D	A
45	A	B	145	A	B	245	C	D	345	A	B	445	A	B	545	B	C	645	A	B
46	A	A	146	E	D	246	A	B	346	C	D	446	A	B	546	B	C	646	A	B
47	A	A	147	A	B	247	A	B	347	A	B	447	A	B	547		C	647	A	B
48	A	B	148	E	D	248	B	C	348	A	B	448	A	B	548	A	A	648	A	B
49	A	B	149	A	B	249	C	D	349	A	B	449	A	B	549	A	B	649	A	B
50	A	A	150	A	B	250	C	D	350	A	B	450	A	B	550	A	D	650	A	D
51	A	B				251	251	D	351	A	B	451	A	B	551	A	B			

52	A	A				252	A	B	352	A	B	452	A	B	552	C	D			
53	A	B				253	A	B	353	A	B	453	A	B	553	A	B			
54						254	A	B	354	A	B	454			554	A	A			
55	A	A				255	B	C	355	C	D	455	A	B	555	A	B			
56	A	B				256	D	C	356	B	C	456	A	B	556	A	B			
57	A	B				257	A	B	357	A	B	457	A	B	557	A	B			
58	A	A				258	A	B	358	B	C	458	A	B	558	C	D			
59	A	A				259	C	D	359	A	B	459	A	B	559	C	D			
60	A	A				260	C	D	360	A	B	460	A	B	560	A	B			
61	A	B	161	B	C	261	A	B	361	C	D	461	A	B	561	A	B			
62	A	B	162	C	D	262	D	B	362	A	A	462	A	B	562	A	B			
63	A	B	163	B	C	263	C	D	363	C	D	463	A	A	563	C	D			
64	A	B	164	B	C	264	A	B	364	A	A	464	A	B	564	C	D			
65	A	B	165	B	C	265	A	A	365	A	B	465	A	B	565	A	D			
66	A	B	166	C	C	266	A	B	366	C	D	466	A	B	566	A	566			
67	A	B	167	B	C	267	A	B	367	A	B	467	A	B	567	C	D			
68	A	B	168			268	A	B	368	A	A	468	A	B	568	A	B			
69	A	B	169	169	C	269	A	B	369	A	B	469	A	B	569	A	B			
70	A	B	170	B	C	270	A	A	370	A	A	470	A	A	570	A	B			
71	A	B				271	A	B	371	B	C	471	A	D	571	A	B			
72	A	B				272	A	B	372	C	D	472	A	B	572	A	A			
73	A	B				273	C	C	373	A	B	473	A	B	573	A	B			
74	A	B				274	A	A	374	C	D	474	A	B	574	A	B			
75	A	B				275	D	C	375	C	D	475	A	B	575	A	B			
76	A	B				276	A	B	376	A	B	476	A	B	576	A	B			
77	A	B				277	A	A	377	A	A	477	A	B	577	A	B			
78	A	B				278	A	A	378	A	A	478	C	D	578	A	B			
79	A	B				279	C	C	379	A	A	479	A	A	579	A	B			
80	A	B				280	A	A	380	C	D	480	A	B	580	B	C			
81	A	B	181	C	D	281		C	381	A	B	481	C	D	581	A	D			
82	A	B	182	***182	D	282	A	A	382	D	C	482	B	C	582	A	A			
83	A	B	183	C		283	A	A	383	A	A	483	A	B	583	A	D			
84	A	A	184			284	A	A	384	A	B	484	A	A	584	A	A			
85	A	B	185			285	A	B	385	A	D	485	A	B	585	D	D			
86	A	B	186	***186	D	286	36	C	386	36	C	486	A	B	586	A	A			
87	A	B	187	***186	D	287	B	C	387	A	B	487	C	D	587	A	A			
88	A	B	188	A	B	288	A	B	388	A	A	488	B	C	588	A	A			
89	A	B	189	A	B	289	A	B	389	A	A	489	489	D	589	A	A			
90	A	90	190			290	A	B	390	D	A	490	C	D	590	A	B			
91	A	B	191	***186	D	291	A	B	391	A	A	491	A	B	591	A	B			
92	A	A	192	A	B	292	C	C	392	A	A	492	C	D	592	D	D			
93	A	B	193	***193	C	293	A	B	393	A	A	493	C	D	593	A	D			
94	A	B	194	A	B	294	B	C	394	A	B	494	A	B	594	A	B			
95	A	B	195	***186	D	295	36	C	395	A	B	495	A	A	595	A	B			
96	A	A	196	***186	D	296	A	A	396	B	C	496	C	D	596	A	A			
97	A	B	197	***186	D	297	A	B	397	397	C	497	B	C	597	A	A			
98	A	B	198	***193	C	298	A	C	398	A	B	498	A	B	598	D	D			
99	A	A	199	***186	D	299	A	A	399	A	B	499	C	D	599	599	A			
100	A	B	200	A	B	300	36	C	400	A	B	500	A	B	600	A	D			

The samples highlighted in red were squenced

The samples highlighted in black were falsed in PCR

The sample highlighted in yellow are from tetroploid cultivars

Appendix 5.3 T-COFFEE multiple sequence alignment of LpSH1

```
T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)
Cedric Notredame
SCORE=1000
*
 BAD AVG GOOD
*
AA-4    : 100
AA-3    : 100
AA-2    : 100
AA-1    : 100
cons    : 100

AA-4    ATEKRORVPSAYNRFIREEIRRIKTNNPDISHREAFSTAAKN
AA-3    ATEKRORVPSAYNRFIREEIORIKTNNPDISHREAFSTAAKN
AA-2    ATEKRRRVPSAYNRFIREEIRRIKTNNPDISHREAFSTAAKN
AA-1    ATEKRQRVPSAYNRFIREEIRRIKTDNPDISHREAFSTAAKN

cons    *****:**************:****:****************
```

Appendix 5.4 T-COFFEE multiple sequence alignment of LpCKX1

```
T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)
Cedric Notredame
CPU TIME:0 sec.
SCORE=1000
*
 BAD AVG GOOD
*
AA-1    : 100
AA-2    : 100
AA-3    : 100
AA-4    : 100
AA-5    : 100
cons    : 100

AA-1    DRGVFGGILQGTDIAGPMVIYPLNKSKWDDSMSAVTPAEDVFYAVSMLFASVANDLKRLQAQNQKILRFCDLAGI
AA-2    DRGVFGGILQGTDIAGPMVIYPLNKSKWDDSMSAVTPAEDVFYAVSMLFSSVANDLKRLQAQNQKILRFCDLAGI
AA-3    DRGVFGGILQGTDIAGPMVIYPLNKSKWDDSMSAVTPAEEVFYAVSMLFASVANDLKRLQAQNQKILRFCDLAGI
AA-4    ----------------MVIYPLNKSKWDDSMSAVTPAEDVFYAVSMLFSSVANDLKRLQAQNQKILRFCDLAGI
AA-5    ----------------MVIYPLNKSKWDDSMSAVTPAEEVFYAVSMLFSSVANDLKRLQAQNQKILRFCDLAGI

cons                    **********************:*********:************************


AA-1    EYKEYLAHYTVRGDWVRHFGSKWDRFVQMKDKYDP----------
AA-2    GYKEYLAHYTVRGDWVRHFGGKWDRFVQMKDKYDP----------
AA-3    GYKEYLAHYTVRGDWVRHFGGKWDRFVQMKDKYDP----------
AA-4    GYKEYLAHYTVRGDWVRHFGGKWDRFVQMKDKYDPKKLLSPGQDI
AA-5    GYKEYLAHYTVRGDWVRHFGGKWDRFVQMKDKYDPKKLLSPGQDI

cons     ******************.**************
```

References

Arthofer, W., Steiner, F. M., & Schlick-Steiner, B. C. (2011). Rapid and cost-effective screening of newly identified microsatellite loci by high-resolution melting analysis. *Molecular Genetics and Genomics, 286*(3-4), 225-235. doi:10.1007/s00438-011-0641-0

Jones, E., Dupal, M., Kölliker, R., Drayton, M., & Forster, J. (2001). Development and characterisation of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.). *Theoretical and Applied Genetics, 102*(2-3), 405-415.

Kayser, M., Kittler, R., Erler, A., Hedman, M., Lee, A. C., Mohyuddin, A., Mehdi, S. Q., Rosser, Z., Stoneking, M., Jobling, M. A., Sajantila, A., & Tyler-Smith, C. (2004). A comprehensive survey of human Y-chromosomal microsatellites. *American Journal of Human Genetics, 74*(6), 1183-1197. doi:10.1086/421531

Kubik, C., Sawkins, M., Meyer, W. A., & Gaut, B. S. (2001). Genetic diversity in seven perennial Ryegrass (*Lolium perenne* L.) cultivars based on SSR markers. *Crop Science, 41*(5), 1565-1572.

Liu, J., Barker, D., Sulc, R., Jang, J., & Wang, G. (2010). Quantifying the proportion of perennial ryegrass cultivars in intra-species mixtures using simple sequence repeat (SSR) and inter-SSR (ISSR) markers and discriminant analysis. *New Zealand Journal of Agricultural Research, 53*(3), 215-226.

Lopez-Giraldez, F., Marmi, J., & Domingo-Roura, X. (2007). High incidence of nonslippage mechanisms generating variability and complexity in *Eurasian badger* microsatellites. *Journal of Heredity, 98*(6), 620-628. doi:10.1093/jhered/esm068

Mackay, J. F., Wright, C. D., & Bonfiglioli, R. G. (2008). A new approach to varietal identification in plants by microsatellite high resolution melting analysis: application to the verification of grapevine and olive cultivars. *Plant Methods, 4*. doi:10.1186/1746-4811-4-8

Mader, E., Lukas, B., & Novak, J. (2008). A strategy to setup codominant microsatellite analysis for High-Resolution-Melting-Curve-analysis (HRM). *Bmc Genetics, 9*. doi:10.1186/1471-2156-9-69

Muleo, R., Colao, M. C., Miano, D., Cirilli, M., Intrieri, M. C., Baldoni, L., & Rugini, E. (2009). Mutation scanning and genotyping by high-resolution DNA melting analysis in olive germplasm. *Genome, 52*(3), 252-260. doi:10.1139/g09-002

Wang, J., Dobrowolski, M. P., Cogan, N. O., Forster, J. W., & Smith, K. F. (2009). Assignment of individual genotypes to specific forage cultivars of perennial ryegrass based on SSR markers. *Crop Science, 49*(1), 49-58.

Wu, S. B., Franks, T. K., Hunt, P., Wirthensohn, M. G., Gibson, J. P., & Sedgley, M. (2010). Discrimination of SNP genotypes associated with complex haplotypes by high resolution melting analysis in almond: implications for improved marker efficiencies. *Molecular Breeding, 25*(2), 351-357. doi:10.1007/s11032-009-9324-5